PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/705, A61K 38/17, C12Q 1/68, G01N 33/68, C07K 16/28

A1

(11) International Publication Number: **WO 98/14571**

(43) International Publication Date: 9 April 1998 (09.04.98)

(21) International Application Number: PCT/US97/17744

(22) International Filing Date: 2 October 1997 (02.10.97)

(30) Priority Data:
08/726,320 3 October 1996 (03.10.96) US

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: AU-YOUNG, Janice [US/US]; 1419 Kains Avenue, Berkeley, CA 94702 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). COLEMAN, Roger [US/US]; 260 Mariposa #2, Mountain View, CA 94041 (US).

(74) Agent: BILLINGS, Lucy, J.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AT, AU, BR, CA, CH, CN, DE, DK, ES, FI, IL, JP, KR, MX, NO, NZ, RU, SE, SG, US, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HUMAN SULFONYLUREA RECEPTOR SURH

(57) Abstract

The present invention provides a human sulfonylurea receptor (SURH) and the polynucleotides which identify and encode SURH. The invention also provides genetically engineered expression vectors and host cells comprising the nucleic acid sequences encoding SURH and methods for producing the protein. The invention also provides pharmaceutical compositions containing SURH, agonists to SURH, or antagonists to SURH, and in the use of such compositions for the prevention or treatment of diseases associated with the expression of SURH. Additionally, the invention provides for the use of antisense molecules to polynucleotides encoding SURH for the treatment of diseases associated with the expression of SURH. The invention also provides diagnostic assays which utilize the polynucleotide, or fragments or the complement thereof, to hybridize to the genomic sequence or transcripts of polynucleotides encoding SURH, or anti-SURH antibodies which specifically bind to SURH.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

HUMAN SULFONYLUREA RECEPTOR SURH

TECHNICAL FIELD

The present invention relates to nucleic acid and amino acid sequences of a novel human sulfonylurea receptor and to the use of these sequences in the diagnosis, study, prevention and treatment of disease.

BACKGROUND ART

ATP-dependent potassium (K_{ATP}) channels serve to couple metabolic state to electrical activity in many types of cells. By hyperpolarizing the cell, K_{ATP} channels limit electrical activity and hence reduce Ca^{2+} entry into muscle and nerve cells. In the pancreas, they are a critical link between blood glucose concentration and insulin secretion.

Sulfonylureas (SUs) are oral hypoglycemics widely used in the treatment of non-insulin dependent diabetes mellitus (NIDDM). SUs stimulate insulin release from pancreatic islet b cells. The mechanism for insulin release involves 1) inhibition of a K_{ATP} channel which sets the b cell resting membrane potential, 2) reduction of K^+ outflow which causes b cell depolarization and 3) the activation of one or more voltage-dependent L-type calcium channels which results in Ca^{2+} influx, exocytosis, and insulin release. SUs such as tolbutamide or glyburide decrease K_{ATP} channel activity, thereby depolarizing the cell and triggering insulin release.

Until recently the K_{ATP} channel and the sulfonylurea receptor (SUR) were thought to be the same molecule (Aguilar-Bryan et al (1995) Science 268:423-426); however, SUR does not possess intrinsic K^+ channel activity (Ammala C et al (1996) Nature 379:545-548). Instead SUR interacts with inward-rectifier K^+ channels, conferring SU and ATP sensitivity to and modulating the activity of these channels (Inagaki N et al (1995) Science 270:1166-1170).

A second isoform of SUR, denoted SUR2, has recently been discovered in rat. This isoform has different tissue distribution and different SU and ATP binding properties from rat SUR (Inagaki N et al (1996) Neuron 16:1011-1017). The channel kinetics of Kir6.2, an inward-rectifier K^+ channel, co-expressed with SUR2 are different than the channel kinetics of Kir6.2 co-expressed with SUR. Based on these observations, it is suggested that a family of structurally related but functionally distinct SURs determine the ATP sensitivity and pharmacological responses of K_{ATP} channels in various tissues (Inagaki N et al (1996), supra).

SURs from rat and hamster consist of 1581 and 1582 amino acids, respectively, with 12 potential membrane-spanning helices (Aguilar-Bryan et al, supra). In addition, the proteins contain two domains having strong similarity to the nucleotide binding folds (NBFs) of the ATP-

binding cassette (ABC) superfamily of proteins. The proposed topology of the rat, hamster, and a recently reported human SUR (GenBank GI 1369844; unpublished) consists of an external amino terminus, nine predicted transmembrane helices, the first cytosolic NBF (NBF-1), four more transmembrane helices, the second cytosolic NBF (NBF-2) and a cytosolic C-terminus. The topology of the SURs are similar to other members of the ABC superfamily including multidrug resistance (MDR) proteins and cystic fibrosis transmembrane regulators (CFTR; Philipson LH and Steiner DF (1995) Science 268:372-373).

The NBFs of ABC superfamily proteins control activity through their interaction with cytosolic nucleotides. In cystic fibrosis, the more frequent and severe disease mutations are located in the nucleotides encoding the two NBFs of the CFTR protein (Tsui L-C (1992) Trends Genet 8:392). Familial persistent hyperinsulinemic hypoglycemia of infancy (PHHI) may be caused by mutations affecting NBF-2 of SUR (Thomas PM et al (1995) Science 268:426-429).

SU-sensitive K_{ATP} channels are present in brain cells and play a role in neurosecretion at nerve terminals. K_{ATP} channels in the substantia nigra, a brain region that shows high SU binding, are inhibited by high glucose concentrations and antidiabetic SUs, and are activated by ATP depletion and anoxia. Furthermore, inhibition of the K_{ATP} channel activates gamma-aminobutyric acid (GABA) release, whereas K_{ATP} channel activation inhibits GABA release (Amoroso S et al (1990) Science 247:852-854; Schmidt-Antomarchi et al (1990) Proc Natl Acad Sci USA 87: 3489-3492).

Action potentials in cardiac cells are modulated by SU compounds binding to SURs. The duration of the action potential of guinea pig cardiac cells was drastically reduced by decreasing intracellular ATP concentrations ($[ATP]_{in}$) by perfusion or by blockade of oxidative phosphorylation. Glibenclamide, an SU compound, was found to restore normal or nearly normal action potentials in these $[ATP]_{in}$-depleted cardiac cells. (Fosset M et al (1988) J Biol Chem 263:7933-7936). Restoration was attributed to inhibition of cardiac K_{ATP} channels by sulfonylurea compounds acting via the SURs.

SURs confer ATP and SU sensitivity to inwardly-rectifying potassium channels, thereby coupling metabolic state to electrical activity in tissues such as brain, pancreas, and heart. SURs are useful in the diagnosis and treatment of diseases related to abnormal K_{ATP} channel function, such as NIDDM and PHHI. The selective modulation of the expression or activities of SURs may allow the successful management of such diseases.

DISCLOSURE OF THE INVENTION

The present invention discloses a human sulfonylurea receptor protein, hereinafter referred to as SURH, having chemical and structural homology to the SUR protein from rat and hamster. Accordingly, the invention features a substantially purified SURH, having the amino acid sequence of SEQ ID NO:1 and the structural characteristics of SURs.

One aspect of the invention features isolated and substantially purified polynucleotides which encode SURH. In a particular aspect, the polynucleotide is the nucleotide sequence of SEQ ID NO:2. In another aspect, the polynucleotide is the nucleotide sequence extending from T_{2780} to A_{2923} of SEQ ID NO:2.

The invention also relates to a polynucleotide sequence comprising the complement of SEQ ID NO:2 or variants thereof. In addition, the invention features nucleotide sequences which hybridize under stringent conditions to SEQ ID NO:2.

The present invention also relates to an expression vector which contains polynucleotides encoding SURH, and the use of said vector to transform or transfect host cells or organisms. The invention also features methods for producing SURH. The present invention also relates to antibodies which bind specifically to SURH polypeptides, and to agonists and antagonists of SURH. The present invention also relates to pharmaceutical compositions comprising SURH, fragments thereof, agonists of SURH, or antagonists of SURH, in conjunction with a suitable pharmaceutical carrier.

BRIEF DESCRIPTION OF DRAWINGS

Figures 1A-1M show the amino acid sequence (SEQ ID NO:1) and the nucleic acid sequence (SEQ ID NO:2) of the human sulfonylurea receptor protein SURH, produced using MacDNAsis software (Hitachi Software Engineering Co Ltd).

Figures 2A-2H show the amino acid sequence alignments among SURH (SEQ ID NO:1), a human SUR isoform (GI 1369844; SEQ ID NO:3), SUR from Norway rat (GI 13115343; SEQ ID NO:4), and SUR from black-bellied hamster (GI 784874; SEQ ID NO:5) produced using the multisequence alignment program of DNAStar software (DNAStar Inc, Madison WI).

MODES FOR CARRYING OUT THE INVENTION

Definitions

"Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand.

Similarly, amino acid sequence as used herein refers to peptide or protein sequence.

"Consensus" as used herein may refer to a nucleic acid sequence 1) which has been resequenced to resolve uncalled bases, 2) which has been extended using XL-PCR (Perkin Elmer, Norwalk CT) in the 5' and/or the 3' direction and resequenced, 3) which has been assembled from the overlapping sequences of more than one Incyte clone GCG Fragment Assembly System, (GCG, Madison WI), or 4) which has been both extended and assembled.

"Peptide nucleic acid" as used herein refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene agents, stop transcript elongation by binding to their complementary (template) strand of nucleic acid (Nielsen PE et al (1993) Anticancer Drug Des 8:53-63).

A "variant" of SURH is defined as an amino acid sequence that is different by one or more amino acid substitutions. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, eg, replacement of leucine with isoleucine. More rarely, a variant may have "nonconservative" changes, eg, replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which and how many amino acid residues may be substituted, inserted or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNAStar software.

A "deletion" is defined as a change in either nucleotide or amino acid sequence in which one or more nucleotides or amino acid residues, respectively, are absent.

An "insertion" or "addition" is that change in a nucleotide or amino acid sequence which has resulted in the addition of one or more nucleotides or amino acid residues, respectively, as compared to the naturally occurring SURH.

A "substitution" results from the replacement of one or more nucleotides or amino acids by different nucleotides or amino acids, respectively.

The term "biologically active" refers to a SURH having structural, regulatory or biochemical functions of the naturally occurring SURH. Likewise, "immunologically active" defines the capability of the natural, recombinant or synthetic SURH, or any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "agonist" refers to a molecule which, when bound to SURH, causes a change in

SURH which modulates the biological activity of SURH. The term "antagonist" refers to a molecule which, when bound to SURH, blocks the binding of an agonist to SURH, which prevents the agonist-induced change in the biological activity of SURH. Agonists and antagonists may include proteins, nucleic acids, carbohydrates, or other molecules which bind to SURH.

The term "modulate" as used herein refers to a change or an alteration in the biological activity of SURH. Modulation may be an increase or a decrease in biological activity, a change in binding characteristics, or any other change in the biological properties of SURH.

The term "derivative" as used herein refers to the chemical modification of a nucleic acid encoding SURH or the encoded SURH. Illustrative of such modifications would be replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative would encode a polypeptide which retains essential biological characteristics of natural SURH.

As used herein, the term "substantially purified" refers to molecules, either nucleic or amino acid sequences, that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated.

"Stringency" typically occurs in a range from about Tm-5°C (5°C below the Tm of the probe) to about 20°C to 25°C below Tm. As will be understood by those of skill in the art, a stringency hybridization can be used to identify or detect identical polynucleotide sequences or to identify or detect similar or related polynucleotide sequences.

The term "hybridization" as used herein shall include "any process by which a strand of nucleic acid joins with a complementary strand through base pairing" (Coombs J (1994) Dictionary of Biotechnology, Stockton Press, New York NY). Amplification as carried out in the polymerase chain reaction technologies is described in Dieffenbach CW and GS Dveksler (1995, PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview NY).

Preferred Embodiments

The invention relates to a human sulfonylurea receptor protein, SURH, initially identified among the cDNAs from a library constructed from human brain tissue (BRAINOT03) and to the use of the nucleic acid and amino acid sequences disclosed herein in the study, diagnosis, prevention and treatment of disease. Northern analysis using the LIFESEQ™ database (Incyte Pharmaceuticals, Palo Alto CA) indicates that SURH-encoding nucleotide sequences are most abundantly transcribed in brain, and are also found in pancreas, breast, uterus and prostate. It

must be noted that naturally occurring expression of SURH is not necessarily limited to these tissues.

The invention also encompasses SURH variants. A preferred SURH variant is one having at least 80% amino acid sequence similarity to the SURH amino acid sequence (SEQ ID NO:1), a more preferred SURH variant is one having at least 90% amino acid sequence similarity to SEQ ID NO:1 and a most preferred SURH variant is one having at least 95% amino acid sequence similarity to SEQ ID NO:1.

The nucleic acid sequence encoding SURH was first identified in the cDNA, Incyte Clone 662342, through a computer-generated search for amino acid sequence alignments. The consensus nucleotide sequence, SEQ ID NO:2, disclosed herein (Figures 1A-1G) encodes the amino acid sequence, SEQ ID NO:1, designated SURH. The consensus nucleotide sequence was extended and assembled from Incyte Clones 1270543 (BRAINOT09); 1332410 (PANCNOT07); 640147 (BRSTNOT03); 641239 (BRSTNOT03); 662342 (BRAINOT03); and 952281 (PANCNOT05) from the LIFESEQ™ database (Incyte Pharmaceuticals, Palo Alto CA).

The present invention is based in part on the chemical homology shown in Figures 2A-2H, among SURH and an SUR isoform from human (GI 1369844, unpublished), and SUR homologs from rat (GI 13115343; Aguilar-Bryan et al. supra) and hamster (GI 784874; Aguilar-Bryan et al. supra). The human isoform, the rat homolog and the hamster homolog have, respectively, 96%, 92% and 90% amino acid sequence identity to SURH.

The SURH protein sequence consists of 1580 amino acids. From the amino acid sequence alignments (Figures 2A-2H) and its hydrophobicity, SURH contains twelve potential membrane-spanning helices located at or near residues 31-51, 75-94, 135-155, 169-189, 306-323, 350-368, 448-469, 540-560, 576-596, 1002-1020, 1063-1076, 1154-1172, and 1276-1296. Furthermore, the protein contains two nucleotide binding fold (NBF) domains encompassing residues 695-893 (NBF-1) and 1356-1534 (NBF-2). The predicted topology of SURH consists of an extracellular amino terminus, nine transmembrane helices, cytosolic NBF-1, four transmembrane helices and cytosolic NBF-2 culminating at the cytosolic C-terminus. In addition, the SURH protein sequence contains 11 potential N-glycosylation sites. Four of these potential N-glycosylation sites, N_{10}, N_{405}, N_{1048} and N_{1058}, reside on predicted extracellular-facing surface loops. As shown is Figures 2A-2H, the SURH amino acid sequence at positions 913 to 923, 925 to 943 and 952 to 960, in the region between NBF-1 and the tenth predicted transmembrane helix, have no identity with the corresponding amino acid residues in a human

SUR isoform (GI 1369844).

The SURH Coding Sequences

The assembled nucleic acid and deduced amino acid sequences of SURH are shown in Figures 1A-1M. In accordance with the invention, any nucleic acid sequence which encodes the amino acid sequence of SURH can be used to generate recombinant molecules which express SURH. In a specific embodiment described herein, a partial sequence encoding SURH was first isolated as Incyte Clone 662342 from a human brain tissue cDNA library (BRAINOT03).

As noted above, there is minimal amino acid sequence identity between a human SUR isoform (GI 1369844) and SURH between amino acids S_{913} to R_{960} of SEQ ID NO:1, which corresponds to nucleotides T_{2780} to A_{2923} of SEQ ID NO:2. Oligonucleotides complementary to this region of SEQ ID NO:2 are highly specific for the SURH of the present invention. Such oligonucleotides are useful for diagnostic and therapeutic applications specific to SURH.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of SURH-encoding nucleotide sequences, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. The invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring SURH, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode SURH and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring SURH under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding SURH or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic expression host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding SURH and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

It is now possible to produce a DNA sequence, or portions thereof, encoding a SURH and its derivatives entirely by synthetic chemistry, after which the synthetic gene may be inserted into

any of the many available DNA vectors and cell systems using reagents that are well known in the art at the time of the filing of this application. Moreover, synthetic chemistry may be used to introduce mutations into a gene encoding SURH.

Also included within the scope of the present invention are polynucleotide sequences that are capable of hybridizing to the nucleotide sequences of Figures 1A-1M under various conditions of stringency. Hybridization conditions are based on the melting temperature (Tm) of the nucleic acid binding complex or probe, as taught in Berger and Kimmel (1987, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol 152, Academic Press, San Diego CA) incorporated herein by reference, and may be used at a defined stringency.

Altered nucleic acid sequences encoding SURH which may be used in accordance with the invention include deletions, insertions or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent SURH. The protein may also show deletions, insertions or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent SURH. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of SURH is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine, valine; glycine, alanine; asparagine, glutamine; serine, threonine phenylalanine, and tyrosine.

Included within the scope of the present invention are alleles of SURH. As used herein, an "allele" or "allelic sequence" is an alternative form of SURH. Alleles result from a mutation, ie, a change in the nucleic acid sequence, and generally produce altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene may have none, one or many allelic forms. Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions or substitutions of amino acids. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

Methods for DNA sequencing are well known in the art and employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase® (US Biochemical Corp, Cleveland OH)), Taq polymerase (Perkin Elmer, Norwalk CT), thermostable T7 polymerase (Amersham, Chicago IL), or combinations of recombinant polymerases and proofreading exonucleases such as the

ELONGASE Amplification System marketed by Gibco BRL (Gaithersburg MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno NV), Peltier Thermal Cycler (PTC200: MJ Research, Watertown MA) and the ABI 377 DNA sequencers (Perkin Elmer).

Extending the Polynucleotide Sequence

The polynucleotide sequence encoding SURH may be extended utilizing partial nucleotide sequence and various methods known in the art to detect upstream sequences such as promoters and regulatory elements. Gobinda et al (1993; PCR Methods Applic 2:318-22) disclose "restriction-site" polymerase chain reaction (PCR) as a direct method which uses universal primers to retrieve unknown sequence adjacent to a known locus. First, genomic DNA is amplified in the presence of primer to a linker sequence and a primer specific to the known region. The amplified sequences are subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR can be used to amplify or extend sequences using divergent primers based on a known region (Triglia T et al (1988) Nucleic Acids Res 16:8186). The primers may be designed using OLIGO® 4.06 Primer Analysis Software (1992; National Biosciences Inc, Plymouth MN), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Capture PCR (Lagerstrom M et al (1991) PCR Methods Applic 1:111-19) is a method for PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA. Capture PCR also requires multiple restriction enzyme digestions and ligations to place an engineered double-stranded sequence into an unknown portion of the DNA molecule before PCR.

Another method which may be used to retrieve unknown sequences is that of Parker JD et al (1991; Nucleic Acids Res 19:3055-60). Additionally, one can use PCR, nested primers and PromoterFinder libraries to walk in genomic DNA (PromoterFinder™ Clontech (Palo Alto CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

Preferred libraries for screening for full length cDNAs are ones that have been

size-selected to include larger cDNAs. Also, random primed libraries are preferred in that they will contain more sequences which contain the 5' and upstream regions of genes. A randomly primed library may be particularly useful if an oligo d(T) library does not yield a full-length cDNA. Genomic libraries are useful for extension into the 5' nontranslated regulatory region.

Capillary electrophoresis may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. Systems for rapid sequencing are available from Perkin Elmer, Beckman Instruments (Fullerton CA), and other companies. Capillary sequencing may employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled devise camera. Output/light intensity is converted to electrical signal using appropriate software (eg. Genotyper™ and Sequence Navigator™ from Perkin Elmer) and the entire process from loading of samples to computer analysis and electronic data display is computer controlled. Capillary electrophoresis is particularly suited to the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample. The reproducible sequencing of up to 350 bp of M13 phage DNA in 30 min has been reported (Ruiz-Martinez MC et al (1993) Anal Chem 65:2851-8).

Expression of the Nucleotide Sequence

In accordance with the present invention, polynucleotide sequences which encode SURH, fragments of the polypeptide, fusion proteins or functional equivalents thereof may be used in recombinant DNA molecules that direct the expression of SURH in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence, may be used to clone and express SURH. As will be understood by those of skill in the art, it may be advantageous to produce SURH-encoding nucleotide sequences possessing non-naturally occurring codons. Codons preferred by a particular prokaryotic or eukaryotic host (Murray E et al (1989) Nuc Acids Res 17:477-508) can be selected, for example, to increase the rate of SURH expression or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, than transcripts produced from naturally occurring sequence.

The nucleotide sequences of the present invention can be engineered in order to alter a coding sequence of SURH for a variety of reasons, including but not limited to, alterations which modify the cloning, processing and/or expression of the gene product. For example, mutations may be introduced using techniques which are well known in the art, eg. site-directed

mutagenesis to insert new restriction sites, to alter glycosylation patterns, to change codon preference, to produce splice variants, etc.

In another embodiment of the invention, a natural, modified or recombinant nucleotide sequence encoding SURH may be ligated to a heterologous sequence to encode a fusion protein. For example, for screening of peptide libraries for inhibitors of SURH activity, it may be useful to encode a chimeric SURH protein that is recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between a SURH sequence and the heterologous protein sequence, so that the SURH may be cleaved and substantially purified away from the heterologous moiety.

In an alternate embodiment of the invention, the coding sequence for SURH may be synthesized, whole or in part, using chemical methods well known in the art (see Caruthers MH et al (1980) Nuc Acids Res Symp Ser 215-23, Horn T et al(1980) Nuc Acids Res Symp Ser 225-32, etc). Alternatively, the protein itself could be produced using chemical methods to synthesize a SURH amino acid sequence, whole or in part. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge JY et al (1995) Science 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer) in accordance with the instructions provided by the manufacturer.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (eg, Creighton (1983) Proteins, Structures and Molecular Principles. WH Freeman and Co, New York NY). The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (eg, the Edman degradation procedure; Creighton, supra). Additionally the amino acid sequence of SURH, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

Expression Systems

In order to express a biologically active SURH, the nucleotide sequence encoding SURH or its functional equivalent is inserted into an appropriate expression vector, ie, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence.

Methods which are well known to those skilled in the art can be used to construct expression vectors containing a SURH coding sequence and appropriate transcriptional or

translational controls. These methods include in vitro recombinant DNA techniques, synthetic techniques and in vivo recombination or genetic recombination. Such techniques are described in Sambrook et al (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY and Ausubel FM et al (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY.

A variety of expression vector/host systems may be utilized to contain and express a SURH coding sequence. These include but are not limited to microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (eg, baculovirus); plant cell systems transfected with virus expression vectors (eg, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with bacterial expression vectors (eg, Ti or pBR322 plasmid); or animal cell systems.

The "control elements" or "regulatory sequences" of these systems vary in their strength and specificities and are those nontranslated regions of the vector, enhancers, promoters, and 3' untranslated regions, which interact with host cellular proteins to carry out transcription and translation. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript® phagemid (Stratagene, LaJolla CA) or pSport1 (Gibco BRL) and ptrp-lac hybrids and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (eg, heat shock, RUBISCO; and storage protein genes) or from plant viruses (eg, viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from the mammalian genes or from mammalian viruses are most appropriate. If it is necessary to generate a cell line that contains multiple copies of SURH, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for SURH. For example, when large quantities of SURH are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be desirable. Such vectors include, but are not limited to, the multifunctional E. coli cloning and expression vectors such as Bluescript® (Stratagene), in which the SURH coding sequence may be ligated into the vector in frame with sequences for the

amino-terminal Met and the subsequent 7 residues of ß-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke & Schuster (1989) J Biol Chem 264:5503-5509); and the like. pGEX vectors (Promega, Madison WI) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems are designed to include heparin, thrombin or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase and PGH may be used. For reviews, see Ausubel et al (supra) and Grant et al (1987) Methods in Enzymology 153:516-544.

In cases where plant expression vectors are used, the expression of a sequence encoding SURH may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV (Brisson et al (1984) Nature 310:511-514) may be used alone or in combination with the omega leader sequence from TMV (Takamatsu et al (1987) EMBO J 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO (Coruzzi et al (1984) EMBO J 3:1671-1680; Broglie et al (1984) Science 224:838-843); or heat shock promoters (Winter J and Sinibaldi RM (1991) Results Probl Cell Differ 17:85-105) may be used. These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. For reviews of such techniques, see Hobbs S or Murry LE in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill New York NY, pp 191-196 or Weissbach and Weissbach (1988) Methods for Plant Molecular Biology, Academic Press, New York NY, pp 421-463.

An alternative expression system which could be used to express SURH is an insect system. In one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. The SURH coding sequence may be cloned into a nonessential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the SURH coding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein coat. The recombinant viruses are then used to infect S. frugiperda cells or Trichoplusia larvae in which SURH is expressed (Smith et al (1983) J Virol 46:584; Engelhard EK et al (1994) Proc Nat Acad Sci 91:3224-7).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, a coding sequence for SURH may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome will result in a viable virus capable of expressing SURH in infected host cells (Logan and Shenk (1984) Proc Natl Acad Sci 81:3655-59). In addition, transcription enhancers, such as the rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be required for efficient translation of a SURH sequence. These signals include the ATG initiation codon and adjacent sequences. In cases where nucleic acid encoding SURH, its initiation codon and upstream sequences are inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon must be provided. Furthermore, the initiation codon must be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate to the cell system in use (Scharf D et al (1994) Results Probl Cell Differ 20:125-62; Bittner et al (1987) Methods in Enzymol 153:516-544).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be important for correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, 293, WI38, etc have specific cellular machinery and characteristic mechanisms for such post-translational activities and may be chosen to ensure the correct modification and processing of the introduced, foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express SURH may be transformed using expression vectors which contain viral origins of replication or endogenous expression elements and a selectable marker gene. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and

recovery of cells which successfully express the introduced sequences. Resistant clumps of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler M et al (1977) Cell 11:223-32) and adenine phosphoribosyltransferase (Lowy I et al (1980) Cell 22:817-23) genes which can be employed in tk- or aprt- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler M et al (1980) Proc Natl Acad Sci 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin F et al (1981) J Mol Biol 150:1-14) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, supra). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman SC and RC Mulligan (1988) Proc Natl Acad Sci 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, ß glucuronidase and its substrate, GUS, and luciferase and its substrate, luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes CA et al (1995) Methods Mol Biol 55:121-131).

Identification of Transformants Containing the Polynucleotide Sequence

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression should be confirmed. For example, if the SURH polynucleotide sequence is inserted within a marker gene sequence, recombinant cells containing SURH can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a SURH sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem SURH as well.

Alternatively, host cells which contain the coding sequence for SURH and express SURH may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridization and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of the nucleic acid or protein.

The presence of the polynucleotide sequence encoding SURH can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes, portions or fragments of SURH-encoding nucleotides. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the SURH sequence to detect transformants containing SURH DNA or RNA. As used herein "oligonucleotides" or "oligomers" refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides which can be used as a probe or amplimer.

A variety of protocols for detecting and measuring the expression of SURH, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on SURH is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton R et al (1990, Serological Methods, a Laboratory Manual, APS Press, St Paul MN) and Maddox DE et al (1983, J Exp Med 158:1211).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to SURH include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the SURH sequence, or any portion of it, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3 or SP6 and labeled nucleotides.

A number of companies such as Pharmacia Biotech (Piscataway NJ), Promega (Madison WI), and US Biochemical Corp (Cleveland OH) supply commercial kits and protocols for these procedures. Suitable reporter molecules or labels include those radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles and the like. Patents teaching the use of such labels include US Patents 3,817,837, 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241. Also, recombinant immunoglobulins may be produced as shown in US Patent No. 4,816,567 incorporated herein by reference.

Purification of SURH

Host cells transformed with a SURH-encoding nucleotide sequence may be cultured under conditions suitable for the expression and recovery of the encoded protein from cell culture. The protein produced by a recombinant cell may be contained intracellularly or secreted depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing SURH can be designed for efficient production and proper transmembrane insertion of SURH into a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may join SURH to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins (Kroll DJ et al (1993) DNA Cell Biol 12:441-53; cf discussion of vectors infra containing fusion proteins).

SURH may also be expressed as a recombinant protein with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle WA). The inclusion of a cleavable linker sequences such as Factor XA or enterokinase (Invitrogen, San Diego CA) between the purification domain and SURH is useful to facilitate purification. One such expression vector provides for expression of a fusion protein compromising an SURH and contains nucleic acid encoding 6 histidine residues followed by thioredoxin and an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography as described in Porath et al (1992) Protein Expression and Purification 3:263-281) while the enterokinase cleavage site provides a means for purifying the protein from the fusion protein.

In addition to recombinant production, fragments of SURH may be produced by direct peptide synthesis using solid-phase techniques (cf Stewart et al (1969) Solid-Phase Peptide Synthesis, WH Freeman Co, San Francisco; Merrifield J (1963) J Am Chem Soc 85:2149-2154). In vitro protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer, Foster City CA) in accordance with the instructions provided by the manufacturer. Various fragments of SURH may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

Uses of SURH

The rationale for the use of polypeptide and polynucleotide sequences disclosed herein is based in part on the chemical and structural homology among SURH and a human isoform, a rat homolog and a hamster homolog of SUR.

The SUR/K_{ATP} channel complex plays a role in neurosecretion and is implicated in the response of the brain to hyper- and hypoglycemia and ischemia. In the pancreas, it is a critical link between blood glucose concentration and insulin secretion. In cardiac cells, action potentials are modulated by sulfonylurea (SU) compounds binding to SURs. Accordingly, SURH may be used in the diagnosis and treatment of diseases and conditions such as, but not limited to, type II diabetes (NIDDM), hyper- and hypoglycemia (including PHHI), cardiac impulse disorders (such as arrhythmias and tachycardias), and other disorders relating to SUR and the SUR/K_{ATP} channel complex.

Some SU therapeutics, while highly effective in the treatment of NIDDM and related disorders, have adverse side-effects. SUs can cause severe and prolonged hypoglycemia, requiring massive glucose infusions and hospitalization for several days. In addition, some SUs can cause skin lesions, including drug-induced erythroderma (exfoliative dermatitis). The isolated SURH protein or its fragments may therefore be useful as a target in drug discovery programs to screen for novel therapeutic molecules with, for example, more desirable binding characteristics, more efficacious metabolic lifetimes, or fewer or less debilitating side-effects than conventional SU therapeutics.

SURH or its fragments may be used to identify other specific molecules with which it binds such as agonists or antagonists.

SURH-specific antibodies are useful for the diagnosis and treatment of conditions and diseases associated with expression of the polypeptides. Antibodies specifically recognizing SURH may be used to quantitate SURH for diagnostic purposes. Therapeutic antibodies may be used to block or modify the interactions between SUs and SURH, or SURH and K_{ATP} channel, in order to treat diseases or conditions associated with SURH and/or the K_{ATP} channel.

In some instances it may be advantageous to suppress SURH expression. Expression of mutant SURH sequences may be suppressed by administration of SURH antisense oligonucleotides.

The SURH nucleic acid sequence of SEQ ID NO:2 can be incorporated into effective eukaryotic expression vectors and directly administered into somatic cells for gene therapy. In

like manner. RNA transcripts produced in vitro may be encapsulated in and administered via liposomes. Such vectors and transcripts may function transiently or may be incorporated into the host chromosomal DNA for longer term expression.

In vivo delivery of genetic constructs into subjects is developed to the point of targeting specific cell types. The delivery to specific cells has been accomplished, for instance, by complexing nucleic acids with proteinous ligands that recognize cell specific receptors which mediate uptake (cf Wu GY et al (1991) J Biol Chem 266:14338-42). Alternatively, recombinant nucleic acid constructs may be injected directly for local uptake and integration (Jiao S et al (1992) Human Gene Therapy 3:21-33).

SURH Antibodies

SURH-specific antibodies are useful for the diagnosis of conditions and diseases associated with expression of SURH. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by a Fab expression library. Neutralizing antibodies, ie, those which inhibit dimer formation, are especially preferred for diagnostics and therapeutics.

SURH for antibody induction does not require biological activity; however, the protein fragment, or oligopeptide must be antigenic. Peptides used to induce specific antibodies may have an amino acid sequence consisting of at least five amino acids, preferably at least 10 amino acids. Preferably, they should mimic a portion of the amino acid sequence of the natural protein and may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of SURH amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule. Procedures well known in the art can be used for the production of antibodies to SURH.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, etc may be immunized by injection with SURH or any portion, fragment or oligopeptide which retains immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are potentially useful human adjuvants.

Monoclonal antibodies to SURH may be prepared using any technique which provides for

the production of antibody molecules by continuous cell lines in culture. These include but are not limited to the hybridoma technique originally described by Koehler and Milstein (1975 Nature 256:495-497), the human B-cell hybridoma technique (Kosbor et al (1983) Immunol Today 4:72; Cote et al (1983) Proc Natl Acad Sci 80:2026-2030) and the EBV-hybridoma technique (Cole et al (1985) Monoclonal Antibodies and Cancer Therapy, Alan R Liss Inc, New York NY, pp 77-96).

In addition, techniques developed for the production of "chimeric antibodies", the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison et al (1984) Proc Natl Acad Sci 81:6851-6855; Neuberger et al (1984) Nature 312:604-608; Takeda et al (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies (US Patent No. 4,946,778) can be adapted to produce SURH-specific single chain antibodies

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in Orlandi et al (1989, Proc Natl Acad Sci 86:3833-3837), and Winter G and Milstein C (1991; Nature 349:293-299).

Antibody fragments which contain specific binding sites for SURH may also be generated. For example, such fragments include, but are not limited to, the F(ab')2 fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse WD et al (1989) Science 256:1275-1281).

A variety of protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the formation of complexes between SURH and its specific antibody and the measurement of complex formation. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two noninterfering epitopes on a specific SURH protein is preferred, but a competitive binding assay may also be employed. These assays are described in Maddox DE et al (1983, J Exp Med 158:1211).

Diagnostic Assays Using SURH Specific Antibodies

Particular SURH antibodies are useful for the diagnosis of conditions or diseases characterized by expression of SURH or in assays to monitor patients being treated with SURH,

agonists or inhibitors. Diagnostic assays for SURH include methods utilizing the antibody and a label to detect SURH in human body fluids or extracts of cells or tissues. The polypeptides and antibodies of the present invention may be used with or without modification. Frequently, the polypeptides and antibodies will be labeled by joining them, either covalently or noncovalently, with a reporter molecule. A wide variety of reporter molecules are known, several of which were described above.

A variety of protocols for measuring SURH, using either polyclonal or monoclonal antibodies specific for the respective protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on SURH is preferred, but a competitive binding assay may be employed. These assays are described, among other places, in Maddox, DE et al (1983, J Exp Med 158:1211).

In order to provide a basis for diagnosis, normal or standard values for SURH expression must be established. This is accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with antibody to SURH under conditions suitable for complex formation which are well known in the art. The amount of standard complex formation may be quantified by comparing various artificial membranes containing known quantities of SURH with both control and disease samples from biopsied tissues. Then, standard values obtained from normal samples may be compared with values obtained from samples from subjects potentially affected by disease. Deviation between standard and subject values establishes the presence of disease state.

Drug Screening

SURH, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening therapeutic compounds in any of a variety of drug screening techniques. The fragment employed in such a test may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between SURH and the agent being tested, may be measured.

Another technique for drug screening which may be used for high throughput screening of compounds having suitable binding affinity to the SURH is described in detail in "Determination of Amino Acid Sequence Antigenicity" by Geysen HN. WO Application 84/03564, published on September 13, 1984, and incorporated herein by reference. In summary, large numbers of

different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with fragments of SURH and washed. Bound SURH is then detected by methods well known in the art. Substantially purified SURH can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding SURH specifically compete with a test compound for binding SURH. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with SURH.

Uses of the Polynucleotide Encoding SURH

A polynucleotide encoding SURH, or any part thereof, may be used for diagnostic and/or therapeutic purposes. For diagnostic purposes, the SURH of this invention may be used to detect and quantitate gene expression in biopsied tissues in which expression of SURH may be implicated. The diagnostic assay is useful to distinguish between absence, presence, and excess expression of SURH and to monitor regulation of SURH levels during therapeutic intervention. Included in the scope of the invention are oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs.

Another aspect of the subject invention is to provide for hybridization or PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding SURH or closely related molecules. The specificity of the probe, whether it is made from a highly specific region, eg, 10 unique nucleotides in the 5' regulatory region, or a less specific region, eg, especially in the 3' region, and the stringency of the hybridization or amplification (maximal, high, intermediate or low) will determine whether the probe identifies only naturally occurring SURH, alleles or related sequences.

Probes may also be used for the detection of related sequences and should preferably contain at least 50% of the nucleotides from any of these SURH encoding sequences. The hybridization probes of the subject invention may be derived from the nucleotide sequence of SEQ ID NO:2 or from genomic sequence including promoter, enhancer elements and introns of the naturally occurring SURH. Hybridization probes may be labeled by a variety of reporter groups, including radionuclides such as ^{32}P or ^{35}S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Other means for producing specific hybridization probes for SURH DNAs include the cloning of nucleic acid sequences encoding SURH or SURH derivatives into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides.

Diagnostic Use

Polynucleotide sequences encoding SURH may be used for the diagnosis of conditions or diseases with which the expression of SURH is associated. For example, polynucleotide sequences encoding SURH may be used in hybridization or PCR assays of fluids or tissues from biopsies to detect SURH expression. The form of such qualitative or quantitative methods may include Southern or northern analysis, dot blot or other membrane-based technologies; PCR technologies; dip stick, pin, chip and ELISA technologies. All of these techniques are well known in the art and are the basis of many commercially available diagnostic kits.

The SURH nucleotide sequence disclosed herein provide the basis for assays that detect activation or induction associated with inflammation or disease. The SURH nucleotide sequence may be labeled by methods known in the art and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After an incubation period, the sample is washed with a compatible fluid which optionally contains a dye (or other label requiring a developer) if the nucleotide has been labeled with an enzyme. After the compatible fluid is rinsed off, the dye is quantitated and compared with a standard. If the amount of dye in the biopsied or extracted sample is significantly elevated over that of a comparable control sample, the nucleotide sequence has hybridized with nucleotide sequences in the sample, and the presence of elevated levels of SURH nucleotide sequences in the sample indicates the presence of the associated inflammation and/or disease.

Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regime in animal studies, in clinical trials, or in monitoring the treatment of an individual patient. In order to provide a basis for the diagnosis of disease, a normal or standard profile for SURH expression must be established. This is accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with SURH, or a portion thereof, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained for normal subjects with a dilution series of SURH run in the

same experiment where a known amount of substantially purified SURH is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients afflicted with SURH-associated diseases. Deviation between standard and subject values is used to establish the presence of disease.

Once disease is established, a therapeutic agent is administered and a treatment profile is generated. Such assays may be repeated on a regular basis to evaluate whether the values in the profile progress toward or return to the normal or standard pattern. Successive treatment profiles may be used to show the efficacy of treatment over a period of several days or several months.

Polymerase Chain Reaction (PCR) as described in US Patent Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the SURH sequence. Such oligomers are generally chemically synthesized, but they may be generated enzymatically or produced from a recombinant source. Oligomers generally comprise two nucleotide sequences, one with sense orientation (5'->3') and one with antisense (3'<-5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

Additionally, methods which may be used to quantitate the expression of a particular molecule include radiolabeling (Melby PC et al 1993 J Immunol Methods 159:235-44) or biotinylating (Duplaa C et al 1993 Anal Biochem 229-36) nucleotides, coamplification of a control nucleic acid, and standard curves onto which the experimental results are interpolated. Quantitation of multiple samples may be speeded up by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation. A definitive diagnosis of this type may allow health professionals to begin aggressive treatment and prevent further worsening of the condition. Similarly, further assays can be used to monitor the progress of a patient during treatment. Furthermore, the nucleotide sequences disclosed herein may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of nucleotide sequences that are currently known such as the triplet genetic code, specific base pair interactions, and the like.

Therapeutic Use

Based upon its homology to the genes encoding SURs and its expression profile, the SURH polynucleotide disclosed herein may provide the basis for the design of molecules for the

treatment of diseases such as NIDDM, PHHI, and other diseases associated with the SUR and/or the K_{ATP} channel.

Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods which are well known to those skilled in the art can be used to construct recombinant vectors which will express antisense SURH. See, for example, the techniques described in Sambrook et al (supra) and Ausubel et al (supra).

The polynucleotides comprising full length cDNA sequence and/or its regulatory elements enable researchers to use SURH as an investigative tool in sense (Youssoufian H and HF Lodish 1993 Mol Cell Biol 13:98-104) or antisense (Eguchi et al (1991) Annu Rev Biochem 60:631-652) regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers, or larger fragments, can be designed from various locations along the coding or control regions.

Genes encoding SURH can be turned off by transfecting a cell or tissue with expression vectors which express high levels of a desired SURH nucleotide fragment. Such constructs can flood cells with untranslatable sense or antisense sequences. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until all copies are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector (Mettler I, personal communication) and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, DNA, RNA or PNA, to the control regions of SURH, ie, the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, eg, between -10 and +10 regions of the leader sequence, are preferred. The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing compromises the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA were reviewed by Gee JE et al (In: Huber BE and BI Carr (1994) Molecular and Immunologic Approaches, Futura Publishing Co. Mt Kisco NY).

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the

ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Within the scope of the invention are engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of RNA encoding SURH.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Antisense molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding SURH. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine and wybutosine as well as acetyl-, methyl-, thio- and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Methods for introducing vectors into cells or tissues include those methods discussed infra and which are equally suitable for in vivo, in vitro and ex vivo therapy. For ex vivo therapy, vectors are introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient is presented in US Patent Nos. 5,399,493 and 5,437,994, disclosed herein by reference. Delivery by transfection and by liposome are quite well known in the art.

Furthermore, the nucleotide sequences for SURH disclosed herein may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including but not limited to such properties as the triplet genetic code and specific base pair interactions.

Detection and Mapping of Related Polynucleotide Sequences

The nucleic acid sequence for SURH can also be used to generate hybridization probes for mapping the naturally occurring genomic sequence. The sequence may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. These include in situ hybridization to chromosomal spreads, flow-sorted chromosomal preparations, or artificial chromosome constructions such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price CM (1993; Blood Rev 7:127-34) and Trask BJ (1991; Trends Genet 7:149-54).

The technique of fluorescent in situ hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY. Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a SURH on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. For example, an STS based map of the human genome was recently published by the Whitehead-MIT Center for Genomic Research (Hudson TJ et al (1995) Science 270:1945-1954). Often the placement of a gene on the chromosome of another mammalian species such as mouse (Whitehead Institute/MIT Center for Genome Research. Genetic Map of the Mouse. Database Release 10. April 28. 1995) may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators

searching for disease genes using positional cloning or other gene discovery techniques. Once a disease or syndrome, such as ataxia telangiectasia (AT), has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti et al (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

Pharmaceutical Compositions

The present invention relates to pharmaceutical compositions which may comprise nucleotides, proteins, antibodies, agonists, antagonists, or inhibitors, alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. Any of these molecules can be administered to a patient alone, or in combination with other agents, drugs or hormones, in pharmaceutical compositions where it is mixed with excipient(s) or pharmaceutically acceptable carriers. In one embodiment of the present invention, the pharmaceutically acceptable carrier is pharmaceutically inert.

Administration of Pharmaceutical Compositions

Administration of pharmaceutical compositions is accomplished orally or parenterally. Methods of parenteral delivery include topical, intra-arterial (directly to the tumor), intramuscular, subcutaneous, intramedullary, intrathecal, intraventricular, intravenous, intraperitoneal, or intranasal administration. In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of "Remington's Pharmaceutical Sciences" (Maack Publishing Co, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active

compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; and gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores are provided with suitable coatings such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, ie, dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders such as lactose or starches, lubricants such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations for parenteral administration include aqueous solutions of active compounds. For injection, the pharmaceutical compositions of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

Manufacture and Storage

The pharmaceutical compositions of the present invention may be manufactured in a manner that known in the art, eg, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents that are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder in 1 mM-50 mM histidine, 0.1%-2% sucrose, 2%-7% mannitol at a pH range of 4.5 to 5.5 that is combined with buffer prior to use.

After pharmaceutical compositions comprising a compound of the invention formulated in a acceptable carrier have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of SURH, such labeling would include amount, frequency and method of administration.

Therapeutically Effective Dose

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, eg, of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model is also used to achieve a desirable concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of protein or its antibodies, antagonists, or inhibitors which ameliorate the symptoms or condition. Therapeutic efficacy and toxicity of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, eg, ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of

dosage for human use. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage is chosen by the individual physician in view of the patient to be treated. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Additional factors which may be taken into account include the severity of the disease state, eg, tumor size and location; age, weight and gender of the patient; diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long acting pharmaceutical compositions might be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature. See US Patent Nos. 4,657,760; 5,206,344; or 5,225,212. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

INDUSTRIAL APPLICABILITY

I cDNA Library Construction

The BRAINOT03 cDNA library was constructed from normal brain tissue removed from a 26 year old male (lot #0003; Mayo Clinic, Rochester MN). The frozen tissue was homogenized and lysed using a Brinkmann Homogenizer Polytron PT-3000 (Brinkmann Instruments, Westbury NJ). The reagents and extraction procedures were used as supplied in the Stratagene RNA Isolation Kit (Cat. # 200345; Stratagene). The lysate was centrifuged over a 5.7 M CsCl cushion using an Beckman SW28 rotor in a Beckman L8-70M Ultracentrifuge (Beckman Instruments) for 18 hours at 25,000 rpm at ambient temperature. The RNA was extracted once with phenol chloroform pH 8.0, once with acid phenol pH 4.0, precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in water and DNase treated for 15 min at 37°C. The RNA was isolated using the Qiagen Oligotex kit (QIAGEN Inc, Chatsworth CA) and used to

construct the cDNA library.

The RNA was handled according to the recommended protocols in the SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning (Cat. #18248-013; Gibco/BRL). cDNAs were fractionated on a Sepharose CL4B column (Cat. #275105, Pharmacia), and those cDNAs exceeding 400 bp were ligated into pSport I. The plasmid pSport I was subsequently transformed into DH5a™ competent cells (Cat. #18258-012, Gibco/BRL).

II Isolation and Sequencing of cDNA Clones

Plasmid DNA was released from the cells and purified using the Miniprep Kit (Cat. # 77468; Advanced Genetic Technologies Corporation, Gaithersburg MD). This kit consists of a 96 well block with reagents for 960 purifications. The recommended protocol was employed except for the following changes: 1) the 96 wells were each filled with only 1 ml of sterile Terrific Broth (Cat. # 22711, LIFE TECHNOLOGIES™, Gaithersburg MD) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) the bacteria were cultured for 24 hours after the wells were inoculated and then lysed with 60 μl of lysis buffer; 3) a centrifugation step employing the Beckman GS-6R @2900 rpm for 5 min was performed before the contents of the block were added to the primary filter plate; and 4) the optional step of adding isopropanol to TRIS buffer was not routinely performed. After the last step in the protocol, samples were transferred to a Beckman 96-well block for storage.

The cDNAs were sequenced by the method of Sanger F and AR Coulson (1975; J Mol Biol 94:441f), using a Hamilton Micro Lab 2200 (Hamilton, Reno NV) in combination with four Peltier Thermal Cyclers (PTC200; MJ Research, Watertown MA) and Applied Biosystems 377 DNA Sequencing Systems (Perkin Elmer), and reading frame was determined.

III Homology Searching of cDNA Clones and Their Deduced Proteins

Each cDNA was compared to sequences in GenBank using a search algorithm developed by Applied Biosystems and incorporated into the INHERIT- 670 Sequence Analysis System. In this algorithm, Pattern Specification Language (TRW Inc, Los Angeles CA) was used to determine regions of homology. The three parameters that determine how the sequence comparisons run were window size, window offset, and error tolerance. Using a combination of these three parameters, the DNA database was searched for sequences containing regions of homology to the query sequence, and the appropriate sequences were scored with an initial value. Subsequently, these homologous regions were examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments were used

to display the results of the homology search.

Peptide and protein sequence homologies were ascertained using the INHERIT™ 670 Sequence Analysis System in a way similar to that used in DNA sequence homologies. Pattern Specification Language and parameter windows were used to search protein databases for sequences containing regions of homology which were scored with an initial value. Dot-matrix homology plots were examined to distinguish regions of significant homology from chance matches.

BLAST, which stands for Basic Local Alignment Search Tool (Altschul SF (1993) J Mol Evol 36:290-300; Altschul, SF et al (1990) J Mol Biol 215:403-10), was used to search for local sequence alignments. BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs. BLAST is useful for matches which do not contain gaps. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP).

An HSP consists of two sequence fragments of arbitrary but equal lengths whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user. The BLAST approach is to look for HSPs between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The parameter E establishes the statistically significant threshold for reporting database sequence matches. E is interpreted as the upper bound of the expected frequency of chance occurrence of an HSP (or set of HSPs) within the context of the entire database search. Any database sequence whose match satisfies E is reported in the program output.

IV Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook et al. supra).

Analogous computer techniques use BLAST (Altschul SF 1993 and 1990, supra) to search for identical or related molecules in nucleotide databases such as GenBank or the LIFESEQ™ database (Incyte, Palo Alto CA). This analysis is much faster than multiple, membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or homologous.

The basis of the search is the product score which is defined as:

$$\frac{\%\text{ sequence identity} \times \%\text{ maximum BLAST score}}{100}$$

and it takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of the search are reported as a list of libraries in which the SURH encoding sequence occurs. Abundance and percentage abundance of the SURH encoding sequence are also reported. Abundance directly reflects the number of times a particular transcript is represented in a cDNA library, and percent abundance is abundance divided by the total number of sequences examined in the cDNA library.

V Extension of SURH to Full Length or to Recover Regulatory Elements

The nucleic acid sequence encoding SURH (SEQ ID NO:2) is used to design oligonucleotide primers for extending a partial nucleotide sequence to full length or for obtaining 5' sequences from genomic libraries. One primer is synthesized to initiate extension in the antisense direction (XLR) and the other is synthesized to extend sequence in the sense direction (XLF). Primers allow the extension of the known SURH nucleotide sequence "outward" generating amplicons containing new, unknown nucleotide sequence for the region of interest (US Patent Application 08/487,112, filed June 7, 1995, specifically incorporated by reference). The initial primers are designed from the cDNA using OLIGO® 4.06 Primer Analysis Software (National Biosciences), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations is avoided.

The original, selected cDNA libraries, or a human genomic library are used to extend the sequence; the latter is most useful to obtain 5' upstream regions. If more extension is necessary or desired, additional sets of primers are designed to further extend the known region.

By following the instructions for the XL-PCR kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix, high fidelity amplification is obtained. Beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, PCR is

performed using the Peltier Thermal Cycler (PTC200: MJ Research. Watertown MA) and the following parameters:

Step 1	94° C for 1 min (initial denaturation)
Step 2	65° C for 1 min
Step 3	68° C for 6 min
Step 4	94° C for 15 sec
Step 5	65° C for 1 min
Step 6	68° C for 7 min
Step 7	Repeat step 4-6 for 15 additional cycles
Step 8	94° C for 15 sec
Step 9	65° C for 1 min
Step 10	68° C for 7:15 min
Step 11	Repeat step 8-10 for 12 cycles
Step 12	72° C for 8 min
Step 13	4° C (and holding)

A 5-10 μl aliquot of the reaction mixture is analyzed by electrophoresis on a low concentration (about 0.6-0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products were selected and cut out of the gel. Further purification involves using a commercial gel extraction method such as QIAQuick™ (QIAGEN Inc). After recovery of the DNA. Klenow enzyme was used to trim single-stranded. nucleotide overhangs creating blunt ends which facilitate religation and cloning.

After ethanol precipitation. the products are redissolved in 13 μl of ligation buffer. 1μl T4-DNA ligase (15 units) and 1μl T4 polynucleotide kinase are added. and the mixture is incubated at room temperature for 2-3 hours or overnight at 16° C. Competent E. coli cells (in 40 μl of appropriate media) are transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium (Sambrook J et al. supra). After incubation for one hour at 37° C. the whole transformation mixture is plated on Luria Bertani (LB)-agar (Sambrook J et al. supra) containing 2xCarb. The following day, several colonies are randomly picked from each plate and cultured in 150 μl of liquid LB/2xCarb medium placed in an individual well of an appropriate. commercially-available. sterile 96-well microtiter plate. The following day, 5 μl of each overnight culture is transferred into a non-sterile 96-well plate and after dilution 1:10 with water. 5 μl of each sample is transferred into a PCR array.

For PCR amplification. 18 μl of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase. a vector primer and one or both of the gene specific primers used for the extension reaction are added to each well. Amplification is performed using the following conditions:

Step 1	94° C for 60 sec
Step 2	94° C for 20 sec
Step 3	55° C for 30 sec
Step 4	72° C for 90 sec
Step 5	Repeat steps 2-4 for an additional 29 cycles
Step 6	72° C for 180 sec
Step 7	4° C (and holding)

Aliquots of the PCR reactions are run on agarose gels together with molecular weight markers. The sizes of the PCR products are compared to the original partial cDNAs, and appropriate clones are selected, ligated into plasmid and sequenced.

VI Labeling and Use of Hybridization Probes

Hybridization probes derived from SEQ ID NO:2 are employed to screen cDNAs, genomic DNAs or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 (National Biosciences), labeled by combining 50 pmol of each oligomer and 250 mCi of [γ-^{32}P] adenosine triphosphate (Amersham, Chicago IL) and T4 polynucleotide kinase (DuPont NEN®, Boston MA). The labeled oligonucleotides are substantially purified with Sephadex G-25 super fine resin column (Pharmacia). A portion containing 10^7 counts per minute of each of the sense and antisense oligonucleotides is used in a typical membrane based hybridization analysis of human genomic DNA digested with one of the following endonucleases (Ase I, Bgl II, Eco RI, Pst I, Xba 1, or Pvu II; DuPont NEN®).

The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR™ film (Kodak, Rochester NY) is exposed to the blots in a Phosphoimager cassette (Molecular Dynamics, Sunnyvale CA) for several hours, hybridization patterns are compared visually.

VII Antisense Molecules

The nucleotide sequence encoding SURH, or any part thereof, is used to inhibit in vivo or in vitro expression of naturally occurring SURH. Although use of antisense oligonucleotides, comprising about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. An oligonucleotide based on the coding sequence of SURH as

shown in Figures 1A-1M is used to inhibit expression of naturally occurring SURH. The complementary oligonucleotide is designed from the most unique 5' sequence as shown in Figures 1A-1M and used either to inhibit transcription by preventing promoter binding to the upstream nontranslated sequence or translation of an SURH transcript by preventing the ribosome from binding. Using an appropriate portion of the leader and 5' sequence of SEQ ID NO:2, an effective antisense oligonucleotide includes any 15-20 nucleotides spanning the region which translates into the signal or early coding sequence of the polypeptide as shown in Figures 1A-1M.

VIII Expression of SURH

Expression of SURH may be accomplished by subcloning the cDNAs into appropriate expression vectors and transfecting the vectors into analogous expression hosts. In this case, the cloning vector , pSport, previously used for the generation of the cDNA library also provides for direct expression of SURH sequences in E. coli. Upstream of the cloning site, this vector contains a promoter for ß-galactosidase, followed by sequence containing the amino-terminal Met and the subsequent 7 residues of ß-galactosidase. Immediately following these eight residues is an engineered bacteriophage promoter useful for artificial priming and transcription and a number of unique restriction sites, including Eco RI, for cloning.

Induction of the isolated, transfected bacterial strain with IPTG using standard methods will produce a fusion protein corresponding to the first seven residues of ß-galactosidase, about 15 residues of "linker", and the peptide encoded within the cDNA. Since cDNA clone inserts are generated by an essentially random process, there is one chance in three that the included cDNA will lie in the correct frame for proper translation. If the cDNA is not in the proper reading frame, it can be obtained by deletion or insertion of the appropriate number of bases by well known methods including in vitro mutagenesis, digestion with exonuclease III or mung bean nuclease, or the inclusion of an oligonucleotide linker of appropriate length.

The SURH cDNA can be shuttled into other vectors known to be useful for expression of protein in specific hosts. Oligonucleotide primers containing cloning sites as well as a segment of DNA (about 25 bases) sufficient to hybridize to stretches at both ends of the target cDNA can be synthesized chemically by standard methods. These primers can then used to amplify the desired gene segment by PCR. The resulting gene segment can be digested with appropriate restriction enzymes under standard conditions and isolated by gel electrophoresis. Alternately, similar gene segments can be produced by digestion of the cDNA with appropriate restriction

enzymes. Using appropriate primers, segments of coding sequence from more than one gene can be ligated together and cloned in appropriate vectors. It is possible to optimize expression by construction of such chimeric sequences.

Suitable expression hosts for such chimeric molecules include, but are not limited to, mammalian cells such as Chinese Hamster Ovary (CHO) and human 293 cells, insect cells such as Sf9 cells, yeast cells such as Saccharomyces cerevisiae, and bacteria such as E. coli. For each of these cell systems, a useful expression vector may also include an origin of replication to allow propagation in bacteria and a selectable marker such as the ß-lactamase antibiotic resistance gene to allow plasmid selection in bacteria. In addition, the vector may include a second selectable marker such as the neomycin phosphotransferase gene to allow selection in transfected eukaryotic host cells. Vectors for use in eukaryotic expression hosts may require RNA processing elements such as 3' polyadenylation sequences if such are not part of the cDNA of interest.

Additionally, the vector may contain promoters or enhancers which increase gene expression. Such promoters are host specific and include MMTV, SV40, and metallothionine promoters for CHO cells; trp, lac, tac and T7 promoters for bacterial hosts; and alpha factor, alcohol oxidase and PGH promoters for yeast. Transcription enhancers, such as the rous sarcoma virus enhancer, may be used in mammalian host cells. Homogeneous cultures of recombinant cells are obtained through standard culture methods. Cellular fractions from cells containing SURH are prepared by solubilization of the whole cell and isolation of subcellular fractions by differential centrifugation, by the addition of detergent, or by other methods well known in the art. These fractions can be used directly in the following assay.

IX SURH Activity

The SU binding activity of SURH or biologically active fragments thereof may be assayed in a competitive binding assay. The competitive binding of 5-[^{125}I]iodo-2-hydroxyglyburide (^{125}I-HGB; Nelson et al (1992) J Biol Chem 267:14928-14933) and an unlabeled SU to SURH is measured by subsequent UV-crosslinking of bound ^{125}I-HGB to the protein. Detergent-solubilized or membrane-bound SURH, or soluble fragments of SURH, are incubated with varying concentrations of unlabeled SU plus a predetermined concentration of ^{125}I-HGB until equilibrium is reached. Aliquots are irradiated at 312 nm to cross-link SURH-bound ^{125}I-HGB to the protein. The irradiated protein samples are electrophoresed on SDS-polyacrylamide gels. The gels are dried and subjected to autoradiography. Bands corresponding

to ^{125}I-labeled SURH are excised from the dried gels, and the radioactivity quantitated in a gamma radiation counter. Data obtained using different concentrations of unlabeled SUs are used to calculate values for the number, affinity, and association of SURH with the candidate SU ligands using an equation such as presented in Nelson et al (supra).

X Production of SURH Specific Antibodies

SURH substantially purified using PAGE electrophoresis (Sambrook, supra) is used to immunize rabbits and to produce antibodies using standard protocols. The amino acid sequence translated from SURH is analyzed using DNAStar software (DNAStar Inc) to determine regions of high immunogenicity and a corresponding oligopolypeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Analysis to select appropriate epitopes, such as those near the C-terminus or in hydrophilic regions is described by Ausubel FM et al (supra).

Typically, the oligopeptides are 15 residues in length, synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry, and coupled to keyhole limpet hemocyanin (KLH, Sigma) by reaction with M-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel FM et al, supra). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for antipeptide activity, for example, by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated, goat anti-rabbit IgG.

XI Purification of Naturally Occurring SURH Using Specific Antibodies

Naturally occurring or recombinant SURH is substantially purified by immunoaffinity chromatography using antibodies specific for SURH. An immunoaffinity column is constructed by covalently coupling SURH antibody to an activated chromatographic resin such as CnBr-activated Sepharose (Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Cellular fractions from cells containing SURH are prepared by solubilization of the whole cell and isolation of subcellular fractions by differential centrifugation, by the addition of detergent, or by other methods well known in the art.

A fractionated SURH-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of SURH (eg, high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/SURH binding (eg, a buffer of pH 2-3 or a high concentration of a

chaotrope such as urea or thiocyanate ion), and SURH is collected.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: INCYTE PHARMACEUTICALS, INC.

(ii) TITLE OF THE INVENTION: NOVEL HUMAN SULFONYLUREA RECEPTOR

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: U.S.
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Filed Herewith

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/726,320
(B) FILING DATE: 02-OCT-1996

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: AF-0001 PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 650-855-0555
(B) TELEFAX: 650-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1580 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Pro Leu Ala Phe Cys Gly Ser Glu Asn His Ser Ala Ala Tyr Arg
 1               5                  10                  15
Val Asp Gln Gly Val Leu Asn Asn Gly Cys Phe Val Asp Ala Leu Asn
            20                  25                  30
```

```
Val Val Pro His Val Phe Leu Leu Phe Ile Thr Phe Pro Ile Leu Phe
        35                  40                  45
Ile Gly Trp Gly Ser Gln Ser Ser Lys Val His Ile His His Ser Thr
    50                  55                  60
Trp Leu His Phe Pro Gly His Asn Leu Arg Trp Ile Leu Thr Phe Met
65                  70                  75                  80
Leu Leu Phe Val Leu Val Cys Glu Ile Ala Glu Gly Ile Leu Ser Asp
                85                  90                  95
Gly Val Thr Glu Ser His His Leu His Leu Tyr Met Pro Ala Gly Met
            100                 105                 110
Ala Phe Met Ala Ala Val Thr Ser Val Val Tyr Tyr His Asn Ile Glu
        115                 120                 125
Thr Ser Asn Phe Pro Lys Leu Leu Ile Ala Leu Leu Val Tyr Trp Thr
    130                 135                 140
Leu Ala Phe Ile Thr Lys Thr Ile Lys Phe Val Lys Phe Leu Asp His
145                 150                 155                 160
Ala Ile Gly Phe Ser Gln Leu Arg Phe Cys Leu Thr Gly Leu Leu Val
                165                 170                 175
Ile Leu Tyr Gly Met Leu Leu Leu Val Glu Val Asn Val Ile Arg Val
            180                 185                 190
Arg Arg Tyr Ile Phe Phe Lys Thr Pro Arg Glu Val Lys Pro Pro Glu
        195                 200                 205
Asp Leu Gln Asp Leu Gly Val Arg Phe Leu Gln Pro Phe Val Asn Leu
    210                 215                 220
Leu Ser Lys Gly Thr Tyr Trp Trp Met Asn Ala Phe Ile Lys Thr Ala
225                 230                 235                 240
His Lys Lys Pro Xaa Asp Leu Arg Ala Ile Gly Lys Leu Pro Ile Ala
                245                 250                 255
Met Arg Ala Leu Thr Asn Tyr Gln Arg Leu Cys Glu Ala Phe Asp Ala
            260                 265                 270
Gln Arg Lys Asp Ile Gln Gly Thr Gln Gly Ala Arg Ala Ile Trp Gln
        275                 280                 285
Ala Leu Ser His Ala Phe Gly Arg Arg Leu Val Leu Ser Ser Thr Phe
    290                 295                 300
Arg Ile Leu Ala Asp Leu Leu Gly Phe Ala Gly Pro Leu Cys Ile Phe
305                 310                 315                 320
Gly Ile Val Asp His Leu Gly Lys Glu Asn Asp Val Phe Gln Pro Lys
                325                 330                 335
Thr Gln Phe Leu Gly Val Tyr Phe Val Ser Ser Gln Glu Phe Leu Ala
            340                 345                 350
Asn Ala Tyr Val Leu Ala Val Leu Leu Phe Leu Ala Leu Leu Leu Gln
        355                 360                 365
Arg Thr Phe Leu Gln Ala Ser Tyr Tyr Val Ala Ile Glu Thr Gly Ile
    370                 375                 380
Asn Leu Arg Gly Ala Ile Gln Thr Lys Ile Tyr Asn Lys Ile Met His
385                 390                 395                 400
Leu Ser Thr Ser Asn Leu Ser Met Gly Glu Met Thr Ala Gly Gln Ile
                405                 410                 415
Cys Asn Leu Val Ala Ile Asp Thr Asn Gln Leu Met Trp Phe Phe Phe
            420                 425                 430
Leu Cys Pro Asn Leu Trp Ala Met Pro Val Gln Ile Ile Val Gly Val
        435                 440                 445
Ile Leu Leu Tyr Tyr Ile Leu Gly Val Ser Ala Leu Ile Gly Ala Ala
    450                 455                 460
Val Ile Ile Leu Leu Ala Pro Val Gln Tyr Phe Val Ala Thr Lys Leu
465                 470                 475                 480
Ser Gln Ala Gln Arg Ser Thr Leu Glu Tyr Ser Asn Glu Arg Leu Lys
                485                 490                 495
Gln Thr Asn Glu Met Leu Arg Gly Ile Lys Leu Leu Lys Leu Tyr Ala
            500                 505                 510
```

```
Trp Glu Asn Ile Phe Arg Thr Arg Val Glu Thr Thr Arg Arg Lys Glu
            515                 520                 525
Met Thr Ser Leu Arg Ala Phe Ala Ile Tyr Thr Ser Ile Ser Ile Phe
    530                 535                 540
Met Asn Thr Ala Ile Pro Ile Ala Ala Val Leu Ile Thr Phe Val Gly
545                 550                 555                 560
His Val Ser Phe Phe Lys Glu Ala Asp Phe Ser Pro Ser Val Ala Phe
                565                 570                 575
Ala Ser Leu Ser Leu Phe His Ile Leu Val Thr Pro Leu Phe Leu Leu
            580                 585                 590
Ser Ser Val Val Arg Ser Thr Val Lys Ala Leu Val Ser Val Gln Lys
        595                 600                 605
Leu Ser Glu Phe Leu Ser Ser Ala Glu Ile Arg Glu Glu Gln Cys Ala
    610                 615                 620
Pro His Glu Pro Thr Pro Gln Gly Pro Ala Ser Lys Tyr Gln Ala Val
625                 630                 635                 640
Pro Leu Arg Val Val Asn Arg Lys Arg Pro Ala Arg Glu Asp Cys Arg
                645                 650                 655
Gly Leu Thr Gly Pro Leu Gln Ser Leu Val Pro Ser Ala Asp Gly Asp
            660                 665                 670
Ala Asp Asn Cys Cys Val Gln Ile Met Gly Gly Tyr Phe Thr Trp Thr
        675                 680                 685
Pro Asp Gly Ile Pro Thr Leu Ser Asn Ile Thr Ile Arg Ile Pro Arg
    690                 695                 700
Gly Gln Leu Thr Met Ile Val Gly Gln Val Gly Cys Gly Lys Ser Ser
705                 710                 715                 720
Leu Leu Leu Ala Ala Leu Gly Glu Met Gln Lys Val Ser Gly Ala Val
                725                 730                 735
Phe Trp Ser Ser Ser Leu Pro Asp Ser Glu Ile Gly Glu Asp Pro Ser
            740                 745                 750
Pro Glu Arg Glu Thr Ala Thr Asp Leu Asp Ile Arg Lys Arg Gly Pro
        755                 760                 765
Val Ala Tyr Ala Ser Gln Lys Pro Trp Leu Leu Asn Ala Thr Val Glu
    770                 775                 780
Glu Asn Ile Ile Phe Glu Ser Pro Phe Asn Lys Gln Arg Tyr Lys Met
785                 790                 795                 800
Val Ile Glu Ala Cys Ser Leu Gln Pro Asp Ile Asp Ile Leu Pro His
                805                 810                 815
Gly Asp Gln Thr Gln Ile Gly Glu Arg Gly Ile Asn Leu Ser Gly Gly
            820                 825                 830
Gln Arg Gln Arg Ile Ser Val Ala Arg Ala Leu Tyr Gln His Ala Asn
        835                 840                 845
Val Val Phe Leu Asp Asp Pro Phe Ser Ala Leu Asp Ile His Leu Ser
    850                 855                 860
Asp His Leu Met Gln Ala Gly Ile Leu Glu Leu Leu Arg Asp Asp Lys
865                 870                 875                 880
Arg Thr Val Val Leu Val Thr His Lys Leu Gln Tyr Leu Pro His Ala
                885                 890                 895
Asp Trp Ile Ile Ala Met Lys Asp Gly Thr Ile Gln Arg Glu Gly Thr
            900                 905                 910
Ser Arg Thr Ser Arg Gly Leu Asn Ala Ser Ser Leu Ser Thr Gly Arg
        915                 920                 925
Pro His Glu Pro Thr Gly Pro Arg Ala Gly Glu Gly Asn Val Thr Glu
    930                 935                 940
Arg Lys Ala Thr Glu Pro Pro Arg Ala Tyr Leu Val Pro Cys Pro Arg
945                 950                 955                 960
Arg Asp Gly Leu Leu Gln Asp Glu Glu Glu Glu Glu Glu Glu Ala Ala
                965                 970                 975
Glu Asn Glu Glu Asp Asp Tyr Leu Ser Ser Met Leu His Gln Arg Ala
            980                 985                 990
```

```
Glu Ile Pro Trp Arg Ala Cys Xaa Lys Tyr Leu Ser Ser Ala Gly Ile
        995                1000                1005
Leu Leu Leu Ser Leu Leu Val Phe Ser Gln Leu Leu Lys His Met Val
   1010                1015                1020
Leu Val Ala Ile Asp Tyr Trp Leu Ala Lys Trp Thr Asp Ser Ala Leu
025                1030                1035                1040
Thr Leu Thr Pro Ala Thr Arg Asn Cys Ser Leu Asn Gln Glu Cys Thr
                1045                1050                1055
Leu Asn Gln Thr Val Tyr Ala Leu Val Phe Thr Val Leu Cys Ser Leu
            1060                1065                1070
Gly Ile Val Leu Cys Leu Val Thr Ser Val Thr Val Glu Trp Thr Gly
        1075                1080                1085
Leu Lys Val Ala Lys Arg Leu His Arg Ser Leu Leu Asn Arg Ile Ile
   1090                1095                1100
Leu Ala Pro Met Arg Phe Phe Glu Thr Thr Pro Leu Gly Ser Ile Leu
105                1110                1115                1120
Asn Arg Phe Ser Ser Asp Cys Asn Thr Ile Asp Gln His Ile Pro Ser
                1125                1130                1135
Thr Leu Glu Cys Leu Ser Arg Ser Thr Leu Leu Cys Val Ser Ala Leu
            1140                1145                1150
Ala Val Ile Ser Tyr Val Thr Pro Val Phe Leu Val Ala Leu Leu Pro
        1155                1160                1165
Leu Ala Ile Val Cys Tyr Phe Ile Gln Lys Tyr Phe Arg Val Ala Ser
   1170                1175                1180
Arg Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu Leu Ser
185                1190                1195                1200
His Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg Ala Phe Arg
                1205                1210                1215
Tyr Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp Ser Asn
            1220                1225                1230
Asn Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu Glu Val
        1235                1240                1245
Arg Met Glu Tyr Ile Gly Ala Cys Val Val Leu Ile Ala Ala Val Thr
   1250                1255                1260
Ser Ile Ser Asn Ser Leu His Arg Glu Leu Ser Ala Gly Leu Val Gly
265                1270                1275                1280
Leu Gly Leu Thr Tyr Ala Leu Met Val Ser Asn Tyr Leu Asn Trp Met
                1285                1290                1295
Val Arg Asn Leu Ala Asp Met Glu Leu Gln Leu Gly Ala Val Lys Arg
            1300                1305                1310
Ile His Gly Leu Leu Lys Thr Glu Ala Glu Ser Tyr Glu Gly Leu Leu
        1315                1320                1325
Ala Pro Ser Leu Ile Pro Lys Asn Trp Pro Asp Gln Gly Lys Ile Gln
   1330                1335                1340
Ile Gln Asn Leu Ser Val Arg Tyr Asp Ser Ser Leu Lys Pro Val Leu
345                1350                1355                1360
Lys His Val Asn Ala Leu Ile Ser Pro Gly Gln Lys Ile Gly Ile Cys
                1365                1370                1375
Gly Arg Thr Gly Ser Gly Lys Ser Ser Phe Ser Leu Ala Phe Phe Arg
            1380                1385                1390
Met Val Asp Thr Phe Glu Gly His Ile Ile Ile Asp Gly Ile Asp Ile
        1395                1400                1405
Ala Lys Leu Pro Leu His Thr Leu Arg Ser Arg Leu Ser Ile Ile Leu
   1410                1415                1420
Gln Asp Pro Val Leu Phe Ser Gly Thr Ile Arg Phe Asn Leu Asp Pro
425                1430                1435                1440
Glu Arg Lys Cys Ser Asp Ser Thr Leu Trp Glu Ala Leu Glu Ile Ala
                1445                1450                1455
Gln Leu Lys Leu Val Val Lys Ala Leu Pro Gly Gly Leu Asp Ala Ile
            1460                1465                1470
```

```
Ile Thr Glu Gly Gly Glu Asn Phe Ser Gln Gly Gln Arg Gln Leu Phe
      1475                1480                1485
Cys Leu Ala Arg Ala Phe Val Arg Lys Thr Ser Ile Phe Ile Met Asp
   1490                1495                1500
Glu Ala Thr Ala Ser Ile Asp Met Ala Thr Glu Asn Ile Leu Gln Lys
505                1510                1515                1520
Val Val Met Thr Ala Phe Ala Asp Arg Thr Val Val Thr Ile Ala His
               1525                1530                1535
Arg Val His Thr Ile Leu Ser Ala Asp Leu Val Ile Val Leu Lys Arg
           1540                1545                1550
Gly Ala Ile Leu Glu Phe Asp Lys Pro Glu Lys Leu Leu Ser Arg Lys
      1555                1560                1565
Asp Ser Val Phe Ala Ser Phe Val Arg Ala Asp Lys
   1570                1575                1580
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4931 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
GAATTCCCGG GTCGACCCAC GCGTCCGCCG CGCCCGCGCC GCCATGCCCC TGGCCTTCTG     60
CGGCAGCGAG AACCACTCGG CCGCCTACCG GGTGGACCAG GGGGTCCTCA ACAACGGCTG    120
CTTTGTGGAC GCCCTCAACG TGGTGCCGCA CGTCTTCCTA CTCTTCATCA CCTTCCCCAT    180
CCTCTTCATT GGATGGGGAA GTCAGAGCTC CAAGGTGCAC ATCCACCACA GCACATGGCT    240
TCATTTCCCT GGCCACAACC TGCGGTGGAT CCTGACCTTC ATGCTGCTCT TCGTCCTGGT    300
GTGTGAGATT GCAGAGGGCA TCCTGTCTGA TGGGGTGACC GAATCCCACC ATCTGCACCT    360
GTACATGCCA GCCGGATGG CGTTCATGGC TGCTGTCACC TCCGTGGTCT ACTATCACAA    420
CATCGAGACT TCCAACTTCC CCAAGCTGCT AATTGCCCTG CTGGTGTATT GGACCCTGGC    480
CTTCATCACC AAGACCATCA AGTTTGTCAA GTTCTTGGAC CACGCCATCG GCTTCTCGCA    540
GCTACGCTTC TGCCTCACAG GGCTGCTGGT GATCCTCTAT GGGATGCTGC TCCTCGTGGA    600
GGTCAATGTC ATCAGGGTGA GGAGATACAT CTTCTTCAAG ACACCGAGGG AGGTGAAGCC    660
TCCCGAGGAC CTGCAAGACC TGGGGGTACG CTTCCTGCAG CCCTTCGTGA ATCTGCTGTC    720
CAAAGGCACC TACTGGTGGA TGAACGCCTT CATCAAGACT GCCCACAAGA AGCCCRTCGA    780
CTTGCGAGCC ATCGGGAAGC TGCCCATCGC CATGAGGGCC CTCACCAACT ACCAACGGCT    840
CTGCGAGGCC TTTGACGCCC AGCGGAAGGA CATTCAGGGC ACTCAAGGTG CCCGGGCCAT    900
CTGGCAGGCA CTCAGCCATG CCTTCGGGAG GCGCCTGGTC CTCAGCAGCA CTTTCCGCAT    960
CTTGGCCGAC CTGCTGGGCT TCGCCGGGCC ACTGTGCATC TTTGGGATCC TGGACCACCT   1020
TGGGAAGGAG AACGACGTCT TCCAGCCCAA GACACAATTT CTCGGGGTTT ACTTTGTCTC   1080
ATCCCAAGAG TTCCTTGCCA ATGCCTACGT CTTAGCTGTG CTTCTGTTCC TTGCCCTCCT   1140
ACTGCAAAGG ACATTTCTGC AAGCATCCTA CTATGTGGCC ATTGAAACTG GAATTAACTT   1200
GAGAGGAGCA ATACAGACCA AGATTTACAA TAAAATTATG CACCTGTCCA CCTCCAACCT   1260
GTCCATGGGA GAAATGACTG CTGGACAGAT CTGYAATCTG GTTGCCATCC ACACCAATCA   1320
GCTCATGTGG TTTTTCTTCT TGTGCCCAAA CCTCTGGGCT ATGCCAGTAC AGATCATTGT   1380
GGGTGTGATT CTCCTCTACT ACATACTCGG AGTCAGTGCC TTAATTGGAG CAGCTGTCAT   1440
CATTCTACTG GCTCCTGTCC AGTACTTCGT GGCCACCAAG CTGTCTCAGG CCCAGCGGAG   1500
CACACTGGAG TATTCCAATG AGCGGCTGAA GCAGACCAAC GAGATGCTCC GCGGCATCAA   1560
GCTGCTGAAG CTGTACGCCT GGGAGAACAT CTTCCGCACG CGGGTGGAGA CGACCCGCAG   1620
GAAGGAGATG ACCAGCCTCA GGGCCTTTGC CATCTATACC TCCATCTCCA TTTTCATGAA   1680
CACGGCCATC CCCATTGCAG CTGTCCTCAT AACTTTCGTG GGCCATGTCA GCTTCTTCAA   1740
```

AGAGGCCGAC TTCTCGCCCT CCGTGGCCTT TGCCTCCCTC TCCCTCTTCC ATATCTTGGT 1800
CACACCGCTG TTCCTGCTGT CCAGTGTGGT CCGATCTACC GTCAAAGCTC TAGTGAGCGT 1860
GCAAAAGCTA AGCGAGTTCC TGTCCAGTGC AGAGATCCGT GAGGAGCAGT GTGCCCCCCA 1920
TGAGCCCACA CCTCAGGGCC CAGCCAGCAA GTACCAGGCG GTGCCCCTCA GGGTTGTGAA 1980
CCGCAAGCGT CCAGCCCGGG AGGATTGTCG GGGCCTCACC GGCCCACTGC AGAGCCTGGT 2040
CCCCAGTGCA GATGGCGATG CTGACAACTG CTGTGTCCAG ATCATGGGAG GCTACTTCAC 2100
GTGGACCCCA GATGGAATCC CCACACTGTC CAACATCACC ATTCGTATCC CCCGAGGCCA 2160
GCTGACTATG ATCGTGGGGC AGGTGGGCTG CGGCAAGTCC TCGCTCCTTC TAGCCGCACT 2220
GGGGGAGATG CAGAAGGTCT CAGGGGCTGT CTTCTGGAGC AGCAGCCTTC CTGACAGCGA 2280
GATAGGAGAG GACCCCAGCC CAGAGCGGGA GACAGCGACC GACTTGGATA TCAGGAAGAG 2340
AGGCCCCGTG GCCTATGCTT CGCAGAAACC ATGGCTGCTA AATGCCACTG TGGAGGAGAA 2400
CATCATCTTT GAGAGTCCCT TCAACAAACA ACGGTACAAG ATGGTCATTG AAGCCTGCTC 2460
TCTGCAGCCA GACATCGACA TCCTGCCCCA TGGAGACCAG ACCCAGATTG GGGAACGGGG 2520
CATCAACCTG TCTGGTGGTC AACGCCAGCG AATCAGTGTG GCCCGAGCCC TCTACCAGCA 2580
CGCCAACGTT GTCTTCTTGG ATGACCCCTT CTCAGCTCTG GATATCCATC TGAGTGACCA 2640
CTTAATGCAG GCCGGCATCC TTGAGCTGCT CCGGGACGAC AAGAGGACAG TGGTCTTAGT 2700
GACCCACAAG CTACAGTACC TGCCCCATGC AGACTGGATC ATTGCCATGA AGGATGGCAC 2760
CATCCAGAGG GAGGGTACCT CAAGGACTTC CAGAGGTCTG AATGCCAGCT CTTTGAGCAC 2820
TGGAAGACCT CATGAACCGA CAGGACCAAG AGCTGGAGAA GGAAATGTCA CAGAGAGAAA 2880
AGCCACAGAG CCACCCAGGG CCTATCTCGT GCCATGTCCT CGAAGGGATG GCCTTCTGCA 2940
GGATGAGGAA GAGGAGGAAG AGGAGGCAGC TGAGAACGAG GAGGATGACT ACCTGTCGTC 3000
CATGCTGCAC CAGCGTGCTG AGATCCCATG GCGAGCCTGC NCCAAGTACC TGTCCTCCGC 3060
CGGCATCCTG CTCCTGTCGT TGCTGGTCTT GTCACAGCTG CTCAAGCACA TGGTCCTGGT 3120
GGCCATCGAC TACTGGCTGG CCAAGTGGAC CGACAGCGCC CTGACCCTGA CCCCTGCAAC 3180
CAGGAACTGC TCCCTCAACC AGGAGTGCAC CCTCAACCAG ACTGTCTATG CCTTGGTGTT 3240
CACGGTGCTC TGCAGCCTGG GCATTGTGCT GTGCCTCGTC ACGTCTGTCA CTGTGGAGTG 3300
GACAGGGCTG AAGGTGGCCA AGAGACTGCA CCGCAGCCTG CTAAACCGGA TCATCCTAGC 3360
CCCCATGAGG TTTTTTGAGA CCACGCCCCT TGGGAGCATC CTGAACAGAT TTTCATCTGA 3420
CTGTAACACC ATCGACCAGC ACATCCCATC CACGCTGGAG TGCCTGAGCC GCTCCACCCT 3480
GCTCTGTGTC TCAGCCCTGG CCGTCATCTC CTATGTCACA CCTGTGTTCC TCGTGGCCCT 3540
CTTGCCCCTG GCCATCGTGT GCTACTTCAT CCAGAAGTAC TTCCGGGTGG CGTCCAGGGA 3600
CCTGCAGCAG CTGGATGACA CCACCCAGCT TCCACTTCTC TCACACTTTG CCGAAACCGT 3660
AGAAGGACTC ACCACCATCC GGGCCTTCAG GTATGAGGCC CGGTTCCAGC AGAAGCTTCT 3720
CGAATACACA GACTCCAACA ACATTGCTTC CCTCTTCCTC ACAGCTGCCA ACAGATGGCT 3780
GGAAGTCCGA ATGGAGTACA TCGGTGCATG TGTGGTGCTC ATCGCAGCGG TGACCTCCAT 3840
CTCCAACTCC CTGCACAGAG AGCTCTCTGC TGGCCTGGTG GGCCTGGGCC TTACCTACGC 3900
CCTAATGGTC TCCAACTACC TCAACTGGAT GGTGAGGAAC CTGGCAGACA TGGAGCTCCA 3960
GCTGGGGGCT GTGAAGCGCA TCCATGGGCT CCTGAAAACC GAGGCAGAGA GCTACGAGGG 4020
GCTCCTGGCA CCATCGCTGA TCCCAAAGAA CTGGCCAGAC CAAGGGAAGA TCCAGATCCA 4080
GAACCTGAGC GTGCGCTACG ACAGCTCCCT GAAGCCGGTG CTGAAGCACG TCAATGCCCT 4140
CATCTCCCCT GGACAGAAGA TCGGGATCTG CGGCCGCACC GGCAGTGGGA AGTCCTCCTT 4200
CTCTCTTGCC TTCTTCCGCA TGGTGGACAC GTTCGAAGGG CACATCATCA TTGATGGCAT 4260
TGACATCGCC AAACTGCCGC TGCACACCCT GCGCTCACGC CTCTCCATCA TCCTGCAGGA 4320
CCCCGTCCTC TTCAGCGGCA CCATCCGATT TAACCTGGAC CCTGAGAGGA AGTGCTCAGA 4380
TAGCACACTG TGGGAGGCCC TGGAAATCGC CCAGCTGAAG CTGGTGGTGA AGGCACTGCC 4440
AGGAGGCCTC GATGCCATCA TCACAGAAGG CGGGGAGAAT TTCAGCCAGG GACAGAGGCA 4500
GCTGTTCTGC CTGGCCCGGG CCTTCGTGAG GAAGACCAGC ATCTTCATCA TGGACGAGGC 4560
CACGGCTTCC ATTGACATGG CCACGGAAAA CATCCTCCAA AAGGTGGTGA TGACAGCCTT 4620
CGCAGACCGC ACTGTGGTCA CCATCGCGCA TCGAGTGCAC ACCATCCTGA GTGCAGACCT 4680
GGTGATCGTC CTGAAGCGGG GTGCCATCCT TGAGTTCGAT AAGCCAGAGA AGCTGCTCAG 4740
CCGGAAGGAC AGCGTCTTCG CCTCCTTCGT CCGTGCAGAC AAGTGACCTG CCAGAGCCCA 4800
AGTGCCATCC CACATTCGGA CCCTGCCCAT ACCCCTGCCT GGGTTTTCTA ACTGTAAATC 4860
ACTTGTAAAT AAATAGATTT GATTATTAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 4920
AAAAAAAAAA A 4931

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1581 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 1369844

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Pro Leu Ala Phe Cys Gly Ser Glu Asn His Ser Ala Ala Tyr Arg
 1               5                  10                  15
Val Asp Gln Gly Val Leu Asn Asn Gly Cys Phe Val Asp Val Leu Asn
            20                  25                  30
Val Val Pro His Val Phe Leu Leu Phe Ile Thr Phe Pro Ile Leu Phe
        35                  40                  45
Ile Gly Trp Gly Ser Gln Ser Ser Lys Val His Ile His His Ser Thr
    50                  55                  60
Trp Leu His Phe Pro Gly His Asn Leu Arg Trp Ile Leu Thr Phe Met
65                  70                  75                  80
Leu Leu Phe Val Leu Val Cys Glu Ile Ala Glu Gly Ile Leu Ser Asp
                85                  90                  95
Gly Val Thr Glu Ser His His Leu His Leu Tyr Met Pro Ala Gly Met
            100                 105                 110
Ala Phe Met Ala Ala Val Thr Ser Val Val Tyr Tyr His Asn Ile Glu
        115                 120                 125
Thr Ser Asn Phe Pro Lys Leu Leu Ile Ala Leu Leu Val Tyr Trp Thr
    130                 135                 140
Leu Ala Phe Ile Thr Lys Thr Ile Lys Phe Val Lys Phe Leu Asp His
145                 150                 155                 160
Ala Ile Ala Phe Ser Gln Val Arg Phe Cys Leu Thr Gly Leu Leu Val
                165                 170                 175
Ile Leu Tyr Gly Met Leu Leu Leu Val Glu Val Asn Val Ile Arg Val
            180                 185                 190
Arg Arg Tyr Ile Phe Phe Lys Thr Pro Arg Glu Val Lys Pro Pro Glu
        195                 200                 205
Asp Leu Gln Asp Leu Gly Val Arg Phe Leu Gln Pro Phe Val Asn Leu
    210                 215                 220
Leu Ser Lys Gly Thr Tyr Trp Trp Met Asn Ala Phe Ile Lys Thr Ala
225                 230                 235                 240
His Lys Lys Pro Ile Asp Leu Arg Ala Ile Gly Lys Leu Pro Ile Ala
                245                 250                 255
Met Arg Ala Leu Thr Asn Tyr Gln Arg Leu Cys Glu Ala Phe Asp Ala
            260                 265                 270
Gln Val Arg Lys Asp Ile Gln Gly Thr Gln Gly Ala Arg Ala Ile Trp
        275                 280                 285
Gln Ala Leu Ser His Ala Phe Gly Arg Arg Leu Val Leu Ser Ser Thr
    290                 295                 300
Phe Arg Ile Leu Ala Asp Leu Leu Gly Phe Ala Gly Pro Leu Cys Ile
305                 310                 315                 320
Phe Gly Ile Val Asp His Leu Gly Lys Glu Asn Asp Val Phe Gln Pro
                325                 330                 335
Lys Thr Gln Phe Leu Gly Val Tyr Phe Val Ser Ser Gln Glu Phe Leu
            340                 345                 350
Ala Asn Ala Tyr Val Leu Ala Val Leu Leu Phe Leu Ala Leu Leu Leu
        355                 360                 365
Gln Arg Thr Phe Leu Gln Ala Ser Tyr Tyr Val Ala Ile Glu Thr Gly
    370                 375                 380
Ile Asn Leu Arg Gly Ala Ile Gln Thr Lys Ile Tyr Asn Lys Ile Met
385                 390                 395                 400
```

His Leu Ser Thr Ser Asn Leu Ser Met Gly Glu Met Thr Ala Gly Gln
405 410 415
Ile Cys Asn Leu Val Ala Ile Asp Thr Asn Gln Leu Met Trp Phe Phe
420 425 430
Phe Leu Cys Pro Asn Leu Trp Ala Met Pro Val Gln Ile Ile Val Gly
435 440 445
Val Ile Leu Leu Tyr Tyr Ile Leu Gly Val Ser Ala Leu Ile Gly Ala
450 455 460
Ala Val Ile Ile Leu Leu Ala Pro Val Gln Tyr Phe Val Ala Thr Lys
465 470 475 480
Leu Ser Gln Ala Gln Arg Thr Thr Leu Glu Tyr Ser Asn Glu Arg Leu
485 490 495
Lys Gln Thr Asn Glu Met Leu Arg Gly Ile Lys Leu Leu Lys Leu Tyr
500 505 510
Ala Trp Glu Asn Ile Phe Arg Thr Arg Val Glu Thr Thr Arg Arg Lys
515 520 525
Glu Met Thr Ser Leu Arg Ala Phe Ala Ile Tyr Thr Ser Ile Ser Ile
530 535 540
Phe Met Asn Thr Ala Ile Pro Ile Ala Ala Val Leu Ile Thr Phe Val
545 550 555 560
Gly His Val Ser Phe Phe Lys Glu Ala Asp Phe Ser Pro Ser Val Ala
565 570 575
Phe Ala Ser Leu Ser Leu Phe His Ile Leu Val Thr Pro Leu Phe Leu
580 585 590
Leu Ser Ser Val Val Arg Ser Thr Val Lys Ala Leu Val Ser Val Gln
595 600 605
Lys Leu Ser Glu Phe Leu Ser Ser Ala Glu Ile Arg Glu Glu Gln Cys
610 615 620
Ala Pro His Glu Pro Thr Pro Gln Gly Pro Ala Ser Lys Tyr Gln Ala
625 630 635 640
Val Pro Leu Arg Val Val Asn Arg Lys Arg Pro Ala Arg Glu Asp Cys
645 650 655
Arg Gly Leu Thr Gly Pro Leu Gln Ser Leu Val Pro Ser Ala Asp Gly
660 665 670
Asp Ala Asp Asn Cys Cys Val Gln Ile Met Gly Gly Tyr Phe Thr Trp
675 680 685
Thr Pro Asp Gly Ile Pro Thr Leu Ser Asn Ile Thr Ile Arg Ile Pro
690 695 700
Arg Gly Gln Leu Thr Met Ile Val Gly Gln Val Gly Cys Gly Lys Ser
705 710 715 720
Ser Leu Leu Leu Ala Ala Leu Gly Glu Met Gln Lys Val Ser Gly Ala
725 730 735
Val Phe Trp Ser Ser Leu Pro Asp Ser Glu Ile Gly Glu Asp Pro Ser
740 745 750
Pro Glu Arg Glu Thr Ala Thr Asp Leu Asp Ile Arg Lys Arg Gly Pro
755 760 765
Val Ala Tyr Ala Ser Gln Lys Pro Trp Leu Leu Asn Ala Thr Val Glu
770 775 780
Glu Asn Ile Ile Phe Glu Ser Pro Phe Asn Lys Gln Arg Tyr Lys Met
785 790 795 800
Val Ile Glu Ala Cys Ser Leu Gln Pro Asp Ile Asp Ile Leu Pro His
805 810 815
Gly Asp Gln Thr Gln Ile Gly Glu Arg Gly Ile Asn Leu Ser Gly Gly
820 825 830
Gln Arg Gln Arg Ile Ser Val Ala Arg Ala Leu Tyr Gln His Ala Asn
835 840 845
Val Val Phe Leu Asp Asp Pro Phe Ser Ala Leu Asp Ile His Leu Ser
850 855 860
Asp His Leu Met Gln Ala Gly Ile Leu Glu Leu Leu Arg Asp Asp Lys
865 870 875 880

```
Arg Thr Val Val Leu Val Thr His Lys Leu Gln Tyr Leu Pro His Ala
                885                 890                 895
Asp Trp Ile Ile Ala Met Lys Asp Gly Thr Ile Gln Arg Glu Gly Thr
            900                 905                 910
Leu Lys Asp Phe Gln Arg Ser Glu Cys Gln Leu Phe Glu His Trp Lys
        915                 920                 925
Thr Leu Met Asn Arg Gln Asp Gln Glu Leu Glu Lys Glu Thr Val Thr
    930                 935                 940
Glu Arg Lys Ala Thr Glu Pro Pro Gln Gly Leu Ser Arg Ala Met Ser
945                 950                 955                 960
Ser Arg Asp Gly Leu Leu Gln Asp Glu Glu Glu Glu Glu Glu Glu Ala
                965                 970                 975
Ala Glu Ser Glu Glu Asp Asp Asn Leu Ser Ser Met Leu His Gln Arg
            980                 985                 990
Ala Glu Ile Pro Trp Arg Ala Cys Ala Lys Tyr Leu Ser Ser Ala Gly
        995                1000                1005
Ile Leu Leu Leu Ser Leu Leu Val Phe Ser Gln Leu Leu Lys His Met
   1010                1015                1020
Val Leu Val Ala Ile Asp Tyr Trp Leu Ala Lys Trp Thr Asp Ser Ala
1025                1030                1035                1040
Leu Thr Leu Thr Pro Ala Ala Arg Asn Cys Ser Leu Ser Gln Glu Cys
                1045                1050                1055
Thr Leu Asp Gln Thr Val Tyr Ala Met Val Phe Thr Ala Val Cys Ser
            1060                1065                1070
Leu Gly Ile Val Leu Cys Leu Val Thr Ser Val Thr Val Glu Trp Thr
        1075                1080                1085
Gly Leu Lys Val Ala Lys Arg Leu His Arg Ser Leu Leu Asn Arg Ile
   1090                1095                1100
Ile Leu Ala Pro Met Arg Phe Phe Glu Thr Thr Pro Leu Gly Ser Ile
1105                1110                1115                1120
Leu Asn Arg Phe Ser Ser Asp Cys Asn Thr Ile Asp Gln His Ile Pro
                1125                1130                1135
Ser Thr Leu Glu Cys Leu Ser Arg Ser Thr Leu Leu Cys Val Ser Ala
            1140                1145                1150
Leu Ala Val Ile Ser Tyr Val Thr Pro Val Phe Leu Val Ala Leu Leu
        1155                1160                1165
Pro Leu Ala Ile Val Cys Tyr Phe Ile Gln Lys Tyr Phe Arg Val Ala
   1170                1175                1180
Ser Arg Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu Leu
1185                1190                1195                1200
Ser His Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg Ala Phe
                1205                1210                1215
Arg Tyr Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp Ser
            1220                1225                1230
Asn Asn Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu Glu
        1235                1240                1245
Val Arg Met Glu Tyr Ile Gly Ala Cys Val Val Leu Ile Ala Ala Val
   1250                1255                1260
Thr Ser Ile Ser Asn Ser Leu His Arg Glu Leu Ser Ala Gly Leu Val
1265                1270                1275                1280
Gly Leu Gly Leu Thr Tyr Ala Leu Met Val Ser Asn Tyr Leu Asn Trp
                1285                1290                1295
Met Val Arg Asn Leu Ala Asp Met Glu Leu Gln Leu Gly Ala Val Lys
            1300                1305                1310
Arg Ile His Gly Leu Leu Lys Thr Glu Ala Glu Ser Tyr Glu Gly Leu
        1315                1320                1325
Leu Ala Pro Ser Leu Ile Pro Lys Asn Trp Pro Asp Gln Gly Lys Ile
   1330                1335                1340
Gln Ile Gln Asn Leu Ser Val Arg Tyr Asp Ser Ser Leu Lys Pro Val
1345                1350                1355                1360
```

```
Leu Lys His Val Asn Ala Leu Ile Ser Pro Gly Gln Lys Ile Gly Ile
                1365                1370                1375
Cys Gly Arg Thr Gly Ser Gly Lys Ser Ser Phe Ser Leu Ala Phe Phe
            1380                1385                1390
Arg Met Val Asp Thr Phe Glu Gly His Ile Ile Ile Asp Gly Ile Asp
        1395                1400                1405
Ile Ala Lys Leu Pro Leu His Thr Leu Arg Ser Arg Leu Ser Ile Ile
    1410                1415                1420
Leu Gln Asp Pro Val Leu Phe Ser Gly Thr Ile Arg Phe Asn Leu Asp
425                 1430                1435                1440
Pro Glu Arg Lys Cys Ser Asp Ser Thr Leu Trp Glu Ala Leu Glu Ile
                1445                1450                1455
Ala Gln Leu Lys Leu Val Val Lys Ala Leu Pro Gly Gly Leu Asp Ala
            1460                1465                1470
Ile Ile Thr Glu Gly Gly Glu Asn Phe Ser Gln Gly Gln Arg Gln Leu
        1475                1480                1485
Phe Cys Leu Ala Arg Ala Phe Val Arg Lys Thr Ser Ile Phe Ile Met
    1490                1495                1500
Asp Glu Ala Thr Ala Ser Ile Asp Met Ala Thr Glu Asn Ile Leu Gln
505                 1510                1515                1520
Lys Val Val Met Thr Ala Phe Ala Asp Arg Thr Val Val Thr Ile Ala
                1525                1530                1535
His Arg Val His Thr Ile Leu Ser Ala Asp Leu Val Ile Val Leu Lys
            1540                1545                1550
Arg Gly Ala Ile Leu Glu Phe Asp Lys Pro Glu Lys Leu Leu Ser Arg
        1555                1560                1565
Lys Asp Ser Val Phe Ala Ser Phe Val Arg Ala Asp Lys
    1570                1575                1580
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1581 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Genbank
(B) CLONE: 1311534

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Pro Leu Ala Phe Cys Gly Thr Glu Asn His Ser Ala Ala Tyr Arg
 1               5                  10                  15
Val Asp Gln Gly Val Leu Asn Asn Gly Cys Phe Val Asp Ala Leu Asn
            20                  25                  30
Val Val Pro His Val Phe Leu Leu Phe Ile Thr Phe Pro Ile Leu Phe
        35                  40                  45
Ile Gly Trp Gly Ser Gln Ser Ser Lys Val His Ile His His Ser Thr
    50                  55                  60
Trp Leu His Phe Pro Gly His Asn Leu Arg Trp Ile Leu Thr Phe Ile
65                  70                  75                  80
Leu Leu Phe Val Leu Val Cys Glu Ile Ala Glu Gly Ile Leu Ser Asp
                85                  90                  95
Gly Val Thr Glu Ser Arg His Leu His Leu Tyr Met Pro Ala Gly Met
            100                 105                 110
```

```
Ala Phe Met Ala Ala Ile Thr Ser Val Val Tyr Tyr His Asn Ile Glu
           115                 120                 125
Thr Ser Asn Phe Pro Lys Leu Leu Ile Ala Leu Leu Ile Tyr Trp Thr
       130                 135                 140
Leu Ala Phe Ile Thr Lys Thr Ile Lys Phe Val Lys Phe Tyr Asp His
145                 150                 155                 160
Ala Ile Gly Phe Ser Gln Leu Arg Phe Cys Leu Thr Gly Leu Leu Val
                165                 170                 175
Ile Leu Tyr Gly Met Leu Leu Leu Val Glu Val Asn Val Ile Arg Val
            180                 185                 190
Arg Arg Tyr Val Phe Phe Lys Thr Pro Arg Glu Val Lys Pro Pro Glu
        195                 200                 205
Asp Leu Gln Asp Leu Gly Val Arg Phe Leu Gln Pro Phe Val Asn Leu
    210                 215                 220
Leu Ser Lys Gly Thr Tyr Trp Trp Met Asn Ala Phe Ile Lys Thr Ala
225                 230                 235                 240
His Lys Lys Pro Ile Asp Leu Arg Ala Ile Gly Lys Leu Pro Ile Ala
                245                 250                 255
Met Arg Ala Leu Thr Asn Tyr Gln Arg Leu Cys Leu Ala Phe Asp Ala
            260                 265                 270
Gln Ala Arg Lys Asp Thr Gln Ser Gln Gln Gly Ala Arg Ala Ile Trp
        275                 280                 285
Arg Ala Leu Cys His Ala Phe Gly Arg Arg Leu Val Leu Ser Ser Thr
    290                 295                 300
Phe Arg Ile Leu Ala Asp Leu Leu Gly Phe Ala Gly Pro Leu Cys Ile
305                 310                 315                 320
Phe Gly Ile Val Asp His Leu Gly Lys Glu Asn His Val Phe Gln Pro
                325                 330                 335
Lys Thr Gln Phe Leu Gly Val Tyr Phe Val Ser Ser Gln Glu Phe Leu
            340                 345                 350
Gly Asn Ala Tyr Val Leu Ala Val Leu Leu Phe Leu Ala Leu Leu Leu
        355                 360                 365
Gln Arg Thr Phe Leu Gln Ala Ser Tyr Tyr Val Ala Ile Glu Thr Gly
    370                 375                 380
Ile Asn Leu Arg Gly Ala Ile Gln Thr Lys Ile Tyr Asn Lys Ile Met
385                 390                 395                 400
His Leu Ser Thr Ser Asn Leu Ser Met Gly Glu Met Thr Ala Gly Gln
                405                 410                 415
Ile Cys Asn Leu Val Ala Ile Asp Thr Asn Gln Leu Met Trp Phe Phe
            420                 425                 430
Phe Leu Cys Pro Asn Leu Trp Ala Met Pro Val Gln Ile Ile Val Gly
        435                 440                 445
Val Ile Leu Leu Tyr Tyr Ile Leu Gly Val Ser Ala Leu Ile Gly Ala
    450                 455                 460
Ala Val Ile Ile Leu Leu Ala Pro Val Gln Tyr Phe Val Ala Thr Lys
465                 470                 475                 480
Leu Ser Gln Ala Gln Arg Thr Thr Leu Glu Tyr Ser Asn Glu Arg Leu
                485                 490                 495
Lys Gln Thr Asn Glu Met Leu Arg Gly Ile Lys Leu Leu Lys Leu Tyr
            500                 505                 510
Ala Trp Glu Asn Ile Phe Cys Ser Arg Val Glu Lys Thr Arg Arg Lys
        515                 520                 525
Glu Met Thr Ser Leu Arg Ala Phe Ala Val Tyr Thr Ser Ile Ser Ile
    530                 535                 540
Phe Met Asn Thr Ala Ile Pro Ile Ala Ala Val Leu Ile Thr Phe Val
545                 550                 555                 560
Gly His Val Ser Phe Phe Lys Glu Ser Asp Phe Ser Pro Ser Val Ala
                565                 570                 575
Phe Ala Ser Leu Ser Leu Phe His Ile Leu Val Thr Pro Leu Phe Leu
            580                 585                 590
```

Leu Ser Ser Val Val Arg Ser Thr Val Lys Ala Leu Val Ser Val Gln
595 600 605
Lys Leu Ser Glu Phe Leu Ser Ser Ala Glu Ile Arg Glu Glu Gln Cys
610 615 620
Ala Pro Arg Glu Pro Ala Pro Gln Gly Gln Ala Gly Lys Tyr Gln Ala
625 630 635 640
Val Pro Leu Lys Val Val Asn Arg Lys Arg Pro Ala Arg Glu Glu Val
645 650 655
Arg Asp Leu Leu Gly Pro Leu Gln Arg Leu Thr Pro Ser Thr Asp Gly
660 665 670
Asp Ala Asp Asn Phe Cys Val Gln Ile Ile Gly Gly Phe Phe Thr Trp
675 680 685
Thr Pro Asp Gly Ile Pro Thr Leu Ser Asn Ile Thr Ile Arg Ile Pro
690 695 700
Arg Gly Gln Leu Thr Met Ile Val Gly Gln Val Gly Cys Gly Lys Ser
705 710 715 720
Ser Leu Leu Leu Ala Thr Leu Gly Glu Met Gln Lys Val Ser Gly Ala
725 730 735
Val Phe Trp Asn Ser Leu Pro Asp Ser Glu Gly Glu Asp Pro Ser Asn
740 745 750
Pro Glu Arg Glu Thr Ala Ala Asp Ser Asp Ala Arg Ser Arg Gly Pro
755 760 765
Val Ala Tyr Ala Ser Gln Lys Pro Trp Leu Leu Asn Ala Thr Val Glu
770 775 780
Glu Asn Ile Thr Phe Glu Ser Pro Phe Asn Lys Gln Arg Tyr Lys Met
785 790 795 800
Val Ile Glu Ala Cys Ser Leu Gln Pro Asp Ile Asp Ile Leu Pro His
805 810 815
Gly Asp Gln Thr Gln Ile Gly Glu Arg Gly Ile Asn Leu Ser Gly Gly
820 825 830
Gln Arg Pro Gly Ile Ser Val Ala Arg Ala Leu Tyr Gln His Thr Asn
835 840 845
Val Val Phe Leu Asp Asp Pro Phe Ser Ala Leu Asp Val His Leu Ser
850 855 860
Asp His Leu Met Gln Ala Gly Ile Leu Glu Leu Leu Arg Asp Asp Lys
865 870 875 880
Arg Thr Val Val Leu Val Thr His Lys Leu Gln Tyr Leu Pro His Ala
885 890 895
Asp Trp Ile Ile Ala Met Lys Asp Gly Thr Ile Gln Arg Glu Gly Thr
900 905 910
Leu Lys Asp Phe Gln Arg Ser Glu Cys Gln Leu Phe Glu His Trp Lys
915 920 925
Thr Leu Met Asn Arg Gln Asp Gln Glu Leu Glu Lys Glu Thr Val Met
930 935 940
Glu Arg Lys Ala Pro Glu Pro Ser Gln Gly Leu Pro Arg Ala Met Ser
945 950 955 960
Ser Arg Asp Gly Leu Leu Leu Asp Glu Asp Glu Glu Glu Glu Glu Ala
965 970 975
Ala Glu Ser Glu Glu Asp Asp Asn Leu Ser Ser Val Leu His Gln Arg
980 985 990
Ala Lys Ile Pro Trp Arg Ala Cys Thr Lys Tyr Leu Ser Ser Ala Gly
995 1000 1005
Ile Leu Leu Leu Ser Leu Leu Val Phe Ser Gln Leu Leu Lys His Met
1010 1015 1020
Val Leu Val Ala Ile Asp Tyr Trp Leu Ala Lys Trp Thr Asp Ser Ala
1025 1030 1035 1040
Leu Val Leu Ser Pro Ala Ala Arg Asn Cys Ser Leu Ser Gln Glu Cys
1045 1050 1055
Ala Leu Asp Gln Ser Val Tyr Ala Met Val Phe Thr Val Leu Cys Ser
1060 1065 1070

```
Leu Gly Ile Ala Leu Cys Leu Val Thr Ser Val Thr Val Glu Trp Thr
        1075                1080                1085
Gly Leu Lys Val Ala Lys Arg Leu His Arg Ser Leu Leu Asn Arg Ile
    1090                1095                1100
Ile Leu Ala Pro Met Arg Phe Phe Glu Thr Thr Pro Leu Gly Ser Ile
105                 1110                1115                1120
Leu Asn Arg Phe Ser Ser Asp Cys Asn Thr Ile Asp Gln His Ile Pro
                1125                1130                1135
Ser Thr Leu Glu Cys Leu Ser Arg Ser Thr Leu Leu Cys Val Ser Ala
            1140                1145                1150
Leu Ala Val Ile Ser Tyr Val Thr Pro Val Phe Leu Val Ala Leu Leu
        1155                1160                1165
Pro Leu Ala Val Val Cys Tyr Phe Ile Gln Lys Tyr Phe Arg Val Ala
    1170                1175                1180
Ser Arg Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu Leu
185                 1190                1195                1200
Ser His Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg Ala Phe
                1205                1210                1215
Arg Tyr Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp Ser
            1220                1225                1230
Asn Asn Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu Glu
        1235                1240                1245
Val Arg Met Glu Tyr Ile Gly Ala Cys Val Val Leu Ile Ala Ala Ala
    1250                1255                1260
Thr Ser Ile Ser Asn Ser Leu His Arg Glu Leu Ser Ala Gly Leu Val
265                 1270                1275                1280
Gly Leu Gly Leu Thr Tyr Ala Leu Met Val Ser Asn Tyr Leu Asn Trp
                1285                1290                1295
Met Val Arg Asn Leu Ala Asp Met Glu Ile Gln Leu Gly Ala Val Lys
            1300                1305                1310
Gly Ile His Thr Leu Leu Lys Thr Glu Ala Glu Ser Tyr Glu Gly Leu
        1315                1320                1325
Leu Ala Pro Ser Leu Ile Pro Lys Asn Trp Pro Asp Gln Gly Lys Ile
    1330                1335                1340
Gln Ile Gln Asn Leu Ser Val Arg Tyr Asp Ser Ser Leu Lys Pro Val
345                 1350                1355                1360
Leu Lys His Val Asn Ala Leu Ile Ser Pro Gly Gln Lys Ile Gly Ile
                1365                1370                1375
Cys Gly Arg Thr Gly Ser Gly Lys Ser Ser Phe Ser Leu Ala Phe Phe
            1380                1385                1390
Arg Met Val Asp Met Phe Glu Gly Arg Ile Ile Ile Asp Gly Ile Asp
        1395                1400                1405
Ile Ala Lys Leu Pro Leu His Thr Leu Arg Ser Arg Leu Ser Ile Ile
    1410                1415                1420
Leu Gln Asp Pro Val Leu Phe Ser Gly Thr Ile Arg Phe Asn Leu Asp
425                 1430                1435                1440
Pro Glu Lys Lys Cys Ser Asp Ser Thr Leu Trp Glu Ala Leu Glu Ile
                1445                1450                1455
Ala Gln Leu Lys Leu Val Val Lys Ala Leu Pro Gly Gly Leu Asp Ala
            1460                1465                1470
Ile Ile Thr Glu Gly Gly Glu Asn Phe Ser Gln Gly Gln Arg Gln Leu
        1475                1480                1485
Phe Cys Leu Ala Arg Ala Phe Val Arg Lys Thr Ser Ile Phe Ile Met
    1490                1495                1500
Asp Glu Ala Thr Ala Ser Ile Asp Met Ala Thr Glu Asn Ile Leu Gln
505                 1510                1515                1520
Lys Val Val Met Thr Ala Phe Ala Asp Arg Thr Val Val Thr Ile Ala
                1525                1530                1535
His Arg Val His Thr Ile Leu Ser Ala Asp Leu Val Met Val Leu Lys
            1540                1545                1550
```

```
Arg Gly Ala Ile Leu Glu Phe Asp Lys Pro Glu Lys Leu Leu Ser Gln
       1555                1560                1565
Lys Asp Ser Val Phe Ala Ser Phe Val Arg Ala Asp Lys
   1570                1575                1580
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1582 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 784874

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Met Pro Leu Ala Phe Cys Gly Thr Glu Asn His Ser Ala Ala Tyr Arg
 1               5                  10                  15
Val Asp Gln Gly Val Leu Asn Asn Gly Cys Phe Val Asp Ala Leu Asn
            20                  25                  30
Val Val Pro His Val Phe Leu Leu Phe Ile Thr Phe Pro Ile Leu Phe
        35                  40                  45
Ile Gly Trp Gly Ser Gln Ser Ser Lys Val His Ile His His Ser Thr
    50                  55                  60
Trp Leu His Phe Pro Gly His Asn Leu Arg Trp Ile Leu Thr Phe Ile
65                  70                  75                  80
Leu Leu Phe Val Leu Val Cys Glu Ile Ala Glu Gly Ile Leu Ser Asp
                85                  90                  95
Gly Val Thr Glu Ser Arg His Leu His Leu Tyr Met Pro Ala Gly Met
            100                 105                 110
Ala Phe Met Ala Ala Ile Thr Ser Val Val Tyr Tyr His Asn Ile Glu
        115                 120                 125
Thr Ser Asn Phe Pro Lys Leu Leu Ile Ala Leu Leu Ile Tyr Trp Thr
    130                 135                 140
Leu Ala Phe Ile Thr Lys Thr Ile Lys Phe Val Lys Phe Tyr Asp His
145                 150                 155                 160
Ala Ile Gly Phe Ser Gln Leu Arg Phe Cys Leu Thr Gly Leu Leu Val
                165                 170                 175
Ile Leu Tyr Gly Met Leu Leu Leu Val Glu Val Asn Val Ile Arg Val
            180                 185                 190
Arg Arg Tyr Ile Phe Phe Lys Thr Pro Arg Glu Val Lys Pro Pro Glu
        195                 200                 205
Asp Leu Gln Asp Leu Gly Val Arg Phe Leu Gln Pro Phe Val Asn Leu
    210                 215                 220
Leu Ser Lys Gly Thr Tyr Trp Trp Met Asn Ala Phe Ile Lys Thr Ala
225                 230                 235                 240
His Lys Lys Pro Ile Asp Leu Arg Ala Ile Ala Lys Leu Pro Ile Ala
                245                 250                 255
Met Arg Ala Leu Thr Asn Tyr Gln Arg Leu Cys Val Ala Phe Asp Ala
            260                 265                 270
Gln Ala Arg Lys Asp Thr Gln Ser Pro Gln Gly Ala Arg Ala Ile Trp
        275                 280                 285
Arg Ala Leu Cys His Ala Phe Gly Arg Arg Leu Ile Leu Ser Ser Thr
    290                 295                 300
Phe Arg Ile Leu Ala Asp Leu Leu Gly Phe Ala Gly Pro Leu Cys Ile
305                 310                 315                 320
```

Phe Gly Ile Val Asp His Leu Gly Lys Glu Asn His Val Phe Gln Pro
325 330 335
Lys Thr Gln Phe Leu Gly Val Tyr Phe Val Ser Ser Gln Glu Phe Leu
340 345 350
Gly Asn Ala Tyr Val Leu Ala Val Leu Leu Phe Leu Ala Leu Leu Leu
355 360 365
Gln Arg Thr Phe Leu Gln Ala Ser Tyr Tyr Val Ala Ile Glu Thr Gly
370 375 380
Ile Asn Leu Arg Gly Ala Ile Gln Thr Lys Ile Tyr Asn Lys Ile Met
385 390 395 400
His Met Ser Thr Ser Asn Leu Ser Met Gly Glu Met Thr Ala Gly Gln
405 410 415
Ile Cys Asn Leu Val Ala Ile Asp Thr Asn Gln Leu Met Trp Phe Phe
420 425 430
Phe Leu Cys Pro Asn Leu Trp Thr Met Pro Val Gln Ile Ile Val Gly
435 440 445
Val Ile Leu Leu Tyr Tyr Ile Leu Gly Val Ser Ala Leu Ile Gly Ala
450 455 460
Ala Val Ile Ile Leu Leu Ala Pro Val Gln Tyr Phe Val Ala Thr Lys
465 470 475 480
Leu Ser Gln Ala Gln Arg Thr Thr Leu Glu His Ser Asn Glu Arg Leu
485 490 495
Lys Gln Thr Asn Glu Met Leu Arg Gly Met Lys Leu Leu Lys Leu Tyr
500 505 510
Ala Trp Glu Ser Ile Phe Cys Ser Arg Val Glu Val Thr Arg Arg Lys
515 520 525
Glu Met Thr Ser Leu Arg Ala Phe Ala Val Tyr Thr Ser Ile Ser Ile
530 535 540
Phe Met Asn Thr Ala Ile Pro Ile Ala Ala Val Leu Ile Thr Phe Val
545 550 555 560
Gly His Val Ser Phe Phe Lys Glu Ser Asp Leu Ser Pro Ser Val Ala
565 570 575
Phe Ala Ser Leu Ser Leu Phe His Ile Leu Val Thr Pro Leu Phe Leu
580 585 590
Leu Ser Ser Val Val Arg Ser Thr Val Lys Ala Leu Val Ser Val Gln
595 600 605
Lys Leu Ser Glu Phe Leu Ser Ser Ala Glu Ile Arg Glu Glu Gln Cys
610 615 620
Ala Pro Arg Glu Pro Ala Pro Gln Gly Gln Ala Gly Lys Tyr Gln Ala
625 630 635 640
Val Pro Leu Lys Val Val Asn Arg Lys Arg Pro Ala Arg Glu Glu Val
645 650 655
Arg Asp Leu Leu Gly Pro Leu Gln Arg Leu Ala Pro Ser Met Asp Gly
660 665 670
Asp Ala Asp Asn Phe Cys Val Gln Ile Ile Gly Gly Phe Phe Thr Trp
675 680 685
Thr Pro Asp Gly Ile Pro Thr Leu Ser Asn Ile Thr Ile Arg Ile Pro
690 695 700
Arg Gly Gln Leu Thr Met Ile Val Gly Gln Val Gly Cys Gly Lys Ser
705 710 715 720
Ser Leu Leu Leu Ala Thr Leu Gly Glu Met Gln Lys Val Ser Gly Ala
725 730 735
Val Phe Trp Asn Ser Asn Leu Pro Asp Ser Glu Gly Arg Gly Pro Gln
740 745 750
Gln Pro Arg Ala Gly Asp Ser Ser Trp Leu Gly Tyr Gln Glu Gln Arg
755 760 765
Pro Arg Gly Tyr Ala Ser Gln Lys Pro Trp Leu Leu Asn Ala Thr Val
770 775 780
Glu Glu Asn Ile Thr Phe Glu Ser Pro Phe Asn Pro Gln Arg Tyr Lys
785 790 795 800

```
Met Val Ile Glu Ala Cys Ser Leu Gln Pro Asp Ile Asp Ile Leu Pro
                805                 810                 815
His Gly Asp Gln Thr Gln Ile Gly Glu Arg Gly Ile Asn Leu Ser Gly
            820                 825                 830
Gly Gln Arg Pro Asp Gln Cys Gly Pro Glu Pro Ser Thr Ser Arg Pro
        835                 840                 845
Met Phe Val Phe Leu Asp Asp Pro Phe Ser Ala Leu Asp Val His Leu
    850                 855                 860
Ser Asp His Leu Met Gln Ala Gly Ile Leu Glu Leu Leu Arg Asp Asp
865                 870                 875                 880
Lys Arg Thr Val Val Leu Val Thr His Lys Leu Gln Tyr Leu Pro His
                885                 890                 895
Ala Asp Trp Ile Ile Ala Met Lys Asp Gly Thr Ile Gln Arg Glu Gly
            900                 905                 910
Thr Leu Lys Asp Phe Gln Arg Ser Glu Cys Gln Leu Phe Glu His Trp
        915                 920                 925
Lys Thr Leu Met Asn Arg Gln Asp Gln Glu Leu Glu Lys Glu Thr Val
    930                 935                 940
Met Glu Arg Lys Ala Ser Glu Pro Ser Gln Gly Leu Pro Arg Ala Met
945                 950                 955                 960
Ser Ser Arg Asp Gly Leu Leu Leu Asp Glu Glu Glu Glu Glu Glu Glu
                965                 970                 975
Ala Ala Glu Ser Glu Glu Asp Asp Asn Leu Ser Ser Val Leu His Gln
            980                 985                 990
Arg Ala Lys Ile Pro Trp Arg Ala Cys Thr Lys Tyr Leu Ser Ser Ala
        995                1000                1005
Gly Ile Leu Leu Leu Ser Leu Leu Val Phe Ser Gln Leu Leu Lys His
   1010                1015                1020
Met Val Leu Val Ala Ile Asp Tyr Trp Leu Ala Lys Trp Thr Asp Ser
1025                1030                1035                1040
Ala Leu Val Leu Ser Pro Ala Ala Arg Asn Cys Ser Leu Ser Gln Glu
                1045                1050                1055
Cys Asp Leu Asp Gln Ser Val Tyr Ala Met Val Phe Thr Leu Leu Cys
            1060                1065                1070
Ser Leu Gly Ile Val Leu Cys Leu Val Thr Ser Val Thr Val Glu Trp
        1075                1080                1085
Thr Gly Leu Lys Val Ala Lys Arg Leu His Arg Ser Leu Leu Asn Arg
   1090                1095                1100
Ile Ile Leu Ala Pro Met Arg Phe Phe Glu Thr Thr Pro Leu Gly Ser
1105                1110                1115                1120
Ile Leu Asn Arg Phe Ser Ser Asp Cys Asn Thr Ile Asp Gln His Ile
                1125                1130                1135
Pro Ser Thr Leu Glu Cys Leu Ser Arg Ser Thr Leu Leu Cys Val Ser
            1140                1145                1150
Ala Leu Thr Val Ile Ser Tyr Val Thr Pro Val Phe Leu Val Ala Leu
        1155                1160                1165
Leu Pro Leu Ala Val Val Cys Tyr Phe Ile Gln Lys Tyr Phe Arg Val
   1170                1175                1180
Ala Ser Arg Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu
1185                1190                1195                1200
Val Ser His Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg Ala
                1205                1210                1215
Phe Arg Tyr Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp
            1220                1225                1230
Ser Asn Asn Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu
        1235                1240                1245
Glu Val Cys Met Glu Tyr Ile Gly Ala Cys Val Val Leu Ile Ala Ala
   1250                1255                1260
Ala Thr Ser Ile Ser Asn Ser Leu His Arg Glu Leu Ser Ala Gly Leu
1265                1270                1275                1280
```

```
Val Gly Leu Gly Leu Thr Tyr Ala Leu Met Val Ser Asn Tyr Leu Asn
                1285                1290                1295
Trp Met Val Arg Asn Leu Ala Asp Met Glu Ile Gln Leu Gly Ala Val
            1300                1305                1310
Lys Arg Ile His Ala Leu Leu Lys Thr Glu Ala Glu Ser Tyr Glu Gly
        1315                1320                1325
Leu Leu Ala Pro Ser Leu Ile Pro Lys Asn Trp Pro Asp Gln Gly Lys
    1330                1335                1340
Ile Gln Ile Gln Asn Leu Ser Val Arg Tyr Asp Ser Ser Leu Lys Pro
1345                1350                1355                1360
Val Leu Lys His Val Asn Thr Leu Ile Ser Pro Gly Gln Lys Ile Gly
                1365                1370                1375
Ile Cys Gly Arg Thr Gly Ser Gly Lys Ser Ser Phe Ser Leu Ala Phe
            1380                1385                1390
Phe Arg Met Val Asp Met Phe Glu Gly Arg Ile Ile Ile Asp Gly Ile
        1395                1400                1405
Asp Ile Ala Lys Leu Pro Leu His Thr Leu Arg Ser Arg Leu Ser Ile
    1410                1415                1420
Ile Leu Gln Asp Pro Val Leu Phe Ser Gly Thr Ile Arg Phe Asn Leu
1425                1430                1435                1440
Asp Pro Glu Lys Lys Cys Ser Asp Ser Thr Leu Trp Glu Ala Leu Glu
                1445                1450                1455
Ile Ala Gln Leu Lys Leu Val Val Lys Ala Leu Pro Gly Gly Leu Asp
            1460                1465                1470
Ala Ile Ile Thr Glu Gly Gly Glu Asn Phe Ser Gln Gly Gln Arg Gln
        1475                1480                1485
Leu Phe Cys Leu Ala Arg Ala Phe Val Arg Lys Thr Ser Ile Phe Ile
    1490                1495                1500
Met Asp Glu Ala Thr Ala Ser Ile Asp Met Ala Thr Glu Asn Ile Leu
1505                1510                1515                1520
Gln Lys Val Val Met Thr Ala Phe Ala Asp Arg Thr Val Val Thr Ile
                1525                1530                1535
Ala His Arg Val His Thr Ile Leu Ser Ala Asp Leu Val Met Val Leu
            1540                1545                1550
Lys Arg Gly Ala Ile Leu Glu Phe Asp Lys Pro Glu Thr Leu Leu Ser
        1555                1560                1565
Gln Lys Asp Ser Val Phe Ala Ser Phe Val Arg Ala Asp Lys
    1570                1575                1580
```

CLAIMS

1. A substantially purified polypeptide comprising the amino acid sequence of SEQ ID NO:1, or fragments thereof.

2. An isolated and purified polynucleotide sequence encoding the polypeptide of claim 1.

3. An isolated polynucleotide sequence comprising the nucleic acid sequence of SEQ ID NO:2 or variants thereof.

4. The polynucleotide sequence of claim 3 wherein said polynucleotide comprises the portion of SEQ ID NO:2 extending from nucleotide T_{2780} to nucleotide A_{2923}.

5. An isolated polynucleotide sequence which is complementary to SEQ ID NO:2 or a variants thereof.

6. An isolated polynucleotide sequence which hybridizes under stringent conditions to SEQ ID NO:2.

7. A hybridization probe comprising SEQ ID NO:2, or fragments thereof.

8. A hybridization probe comprising the polynucleotide sequence of claim 4.

9. A recombinant expression vector containing the polynucleotide sequence of claim 3.

10. A recombinant host cell containing the expression vector of claim 9.

11. A method for producing the polypeptide of SEQ ID NO:1, or fragments thereof, the method comprising the steps of:

a) culturing the host cell of claim 10 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

12. A pharmaceutical composition comprising the polypeptide of SEQ ID NO:1 in conjunction with a pharmaceutically acceptable excipient.

13. A purified antibody which binds specifically to the polypeptide of claim 1.

14. A purified agonist which specifically modulates the activity of the polypeptide of claim 1.

15. A purified antagonist which specifically modulates the activity of the polypeptide of claim 1.

16. A method for the detection of polynucleotides encoding sulfonylurea receptor protein in a biological sample comprising the steps of:

a) hybridizing the probe of claim 7 to nucleic acid material, thereby forming a

hybridization complex. and

b) detecting said hybridization complex. wherein the presence of said complex correlates with the presence of a polynucleotide encoding sulfonylurea receptor protein in said biological sample.

17. The method of claim 16 wherein before hybridization. the nucleic acid material of the biological sample is amplified by the polymerase chain reaction.

```
                9          18          27          36          45          54
5'   G AAT TCC CGG GTC GAC CCA CGC GTC CGC CGC GCC CGC GCC GCC ATG CCC CTG
                                                              M   P   L

               63          72          81          90          99         108
GCC TTC TGC GGC AGC GAG AAC CAC TCG GCC GCC TAC CGG GTG GAC CAG GGG GTC
A   F   C   G   S   E   N   H   S   A   A   Y   R   V   D   Q   G   V

              117         126         135         144         153         162
CTC AAC AAC GGC TGC TTT GTG GAC GCG CTC AAC GTG GTG CCG CAC GTC TTC CTA
L   N   N   G   C   F   V   D   A   L   N   V   V   P   H   V   F   L

              171         180         189         198         207         216
CTC TTC ATC ACC TTC CCC ATC CTC TTC ATT GGA TGG GGA AGT CAG AGC TCC AAG
L   F   I   T   F   P   I   L   F   I   G   W   G   S   Q   S   S   K

              225         234         243         252         261         270
GTG CAC ATC CAC CAC AGC ACA TGG CTT CAT TTC CCT GGG CAC AAC CTG CGG TGG
V   H   I   H   H   S   T   W   L   H   F   P   G   H   N   L   R   W

              279         288         297         306         315         324
ATC CTG ACC TTC ATG CTG CTC TTC GTC CTG GTG TGT GAG ATT GCA GAG GGC ATC
I   L   T   F   M   L   L   F   V   L   V   C   E   I   A   E   G   I

              333         342         351         360         369         378
CTG TCT GAT GGG GTG ACC GAA TCC CAC CAT CTG CAC CTG TAC ATG CCA GCC GGG
L   S   D   G   V   T   E   S   H   H   L   H   L   Y   M   P   A   G
```

FIGURE 1A

```
          387                 396                 405                 414                 423                 432
ATG GCG TTC ATG GCT GCT GTC ACC TCC GTG GTC TAC TAT CAC AAC ATC GAG ACT
M   A   F   M   A   A   V   T   S   V   V   Y   Y   H   N   I   E   T

          441                 450                 459                 468                 477                 486
TCC AAC TTC CCC AAG CTG CTA ATT GCC CTG CTG GTG TAT TGG ACC CTG GCC TTC
S   N   F   P   K   L   L   I   A   L   L   V   Y   W   T   L   A   F

          495                 504                 513                 522                 531                 540
ATC ACC AAG ACC ATC AAG TTT GTC AAG TTC TTG GAC CAC GCC ATC GGC TTC TCG
I   T   K   T   I   K   F   V   K   F   L   D   H   A   I   G   F   S

          549                 558                 567                 576                 585                 594
CAG CTA CGC TTC TGC CTC ACA GGG CTG CTG GTG ATC CTC TAT GGG ATG CTG CTC
Q   L   R   F   C   L   T   G   L   L   V   I   L   Y   G   M   L   L

          603                 612                 621                 630                 639                 648
CTC GTG GAG GTC AAT GTC ATC AGG GTG AGG AGA TAC ATC TTC TTC AAG ACA CCG
L   V   E   V   N   V   I   R   V   R   R   Y   I   F   F   K   T   P

          657                 666                 675                 684                 693                 702
AGG GAG GTG AAG CCT CCC GAG GAC CTG CAA GAC CTG GGG GTA CGC TTC CTG CAG
R   E   V   K   P   P   E   D   L   Q   D   L   G   V   R   F   L   Q

          711                 720                 729                 738                 747                 756
CCC TTC GTG AAT CTG CTG TCC AAA GGC ACC TAC TGG TGG ATG AAC GCC TTC ATC
P   F   V   N   L   L   S   K   G   T   Y   W   W   M   N   A   F   I
```

FIGURE 1B

```
          765        774        783        792        801        810
AAG ACT GCC CAC AAG AAG CCC RTC GAC TTG CGA GCC ATC GGG AAG CTG CCC ATC
K   T   A   H   K   K   P   X   D   L   R   A   I   G   K   L   P   I

          819        828        837        846        855        864
GCC ATG AGG GCC CTC ACC AAC TAC CAA CGG CTC TGC GAG GCC TTT GAC GCC CAG
A   M   R   A   L   T   N   Y   Q   R   L   C   E   A   F   D   A   Q

          873        882        891        900        909        918
CGG AAG GAC ATT CAG GGC ACT CAA GGT GCC CGG GCC ATC TGG CAG GCA CTC AGC
R   K   D   I   Q   G   T   Q   G   A   R   A   I   W   Q   A   L   S

          927        936        945        954        963        972
CAT GCC TTC GGG AGG CGC CTG GTC CTC AGC AGC ACT TTC CGC ATC TTG GCC GAC
H   A   F   G   R   R   L   V   L   S   S   T   F   R   I   L   A   D

          981        990        999       1008       1017       1026
CTG CTG GGC TTC GCC GGG CCA CTG TGC ATC TTT GGG ATC GTG GAC CAC CTT GGG
L   L   G   F   A   G   P   L   C   I   F   G   I   V   D   H   L   G

         1035       1044       1053       1062       1071       1080
AAG GAG AAC GAC GTC TTC CAG CCC AAG ACA CAA TTT CTC GGG GTT TAC TTT GTC
K   E   N   D   V   F   Q   P   K   T   Q   F   L   G   V   Y   F   V

         1089       1098       1107       1116       1125       1134
TCA TCC CAA GAG TTC CTT GCC AAT GCC TAC GTC TTA GCT GTG CTT CTG TTC CTT
S   S   Q   E   F   L   A   N   A   Y   V   L   A   V   L   L   F   L
```

FIGURE 1C

```
        1143        1152        1161        1170        1179        1188
GCC CTC CTA CTG CAA AGG ACA TTT CTG CAA GCA TCC TAC TAT GTG GCC ATT GAA
A   L   L   L   Q   R   T   F   L   Q   A   S   Y   Y   V   A   I   E

        1197        1206        1215        1224        1233        1242
ACT GGA ATT AAC TTG AGA GGA GCA ATA CAG ACC AAG ATT TAC AAT AAA ATT ATG
T   G   I   N   L   R   G   A   I   Q   T   K   I   Y   N   K   I   M

        1251        1260        1269        1278        1287        1296
CAC CTG TCC ACC TCC AAC CTG TCC ATG GGA GAA ATG ACT GCT GGA CAG ATC TGY
H   L   S   T   S   N   L   S   M   G   E   M   T   A   G   Q   I   C

        1305        1314        1323        1332        1341        1350
AAT CTG GTT GCC ATC GAC ACC AAT CAG CTC ATG TGG TTT TTC TTC TTG TGC CCA
N   L   V   A   I   D   T   N   Q   L   M   W   F   F   F   L   C   P

        1359        1368        1377        1386        1395        1404
AAC CTC TGG GCT ATG CCA GTA CAG ATC ATT GTG GGT GTG ATT CTC CTC TAC TAC
N   L   W   A   M   P   V   Q   I   I   V   G   V   I   L   L   Y   Y

        1413        1422        1431        1440        1449        1458
ATA CTC GGA GTC AGT GCC TTA ATT GGA GCA GCT GTC ATC ATT CTA CTG GCT CCT
I   L   G   V   S   A   L   I   G   A   A   V   I   I   L   L   A   P

        1467        1476        1485        1494        1503        1512
GTC CAG TAC TTC GTG GCC ACC AAG CTG TCT CAG GCC CAG CGG AGC ACA CTG GAG
V   Q   Y   F   V   A   T   K   L   S   Q   A   Q   R   S   T   L   E
```

FIGURE 1D

```
   1521        1530        1539        1548        1557        1566
TAT TCC AAT GAG CGG CTG AAG CAG ACC AAC GAG ATG CTC CGC GGC ATC AAG CTG
Y   S   N   E   R   L   K   Q   T   N   E   M   L   R   G   I   K   L

       1575        1584        1593        1602        1611        1620
CTG AAG CTG TAC GCC TGG GAG AAC ATC TTC CGC ACG CGG GTG GAG ACG ACC CGC
L   K   L   Y   A   W   E   N   I   F   R   T   R   V   E   T   T   R

       1629        1638        1647        1656        1665        1674
AGG AAG GAG ATG ACC AGC CTC AGG GCC TTT GCC ATC TAT ACC TCC ATC TCC ATT
R   K   E   M   T   S   L   R   A   F   A   I   Y   T   S   I   S   I

       1683        1692        1701        1710        1719        1728
TTC ATG AAC ACG GCC ATC CCC ATT GCA GCT GTC CTC ATA ACT TTC GTG GGC CAT
F   M   N   T   A   I   P   I   A   A   V   L   I   T   F   V   G   H

       1737        1746        1755        1764        1773        1782
GTC AGC TTC TTC AAA GAG GCC GAC TTC TCG CCC TCC GTG GCC TTT GCC TCC CTC
V   S   F   F   K   E   A   D   F   S   P   S   V   A   F   A   S   L

       1791        1800        1809        1818        1827        1836
TCC CTC TTC CAT ATC TTG GTC ACA CCG CTG TTC CTG CTG TCC AGT GTG GTC CGA
S   L   F   H   I   L   V   T   P   L   F   L   L   S   S   V   V   R

       1845        1854        1863        1872        1881        1890
TCT ACC GTC AAA GCT CTA GTG AGC GTG CAA AAG CTA AGC GAG TTC CTG TCC AGT
S   T   V   K   A   L   V   S   V   Q   K   L   S   E   F   L   S   S
```

FIGURE 1E

```
      1899       1908       1917       1926       1935       1944
GCA GAG ATC CGT GAG GAG CAG TGT GCC CCC CAT GAG CCC ACA CCT CAG GGC CCA
A   E   I   R   E   E   Q   C   A   P   H   E   P   T   P   Q   G   P

      1953       1962       1971       1980       1989       1998
GCC AGC AAG TAC CAG GCG GTG CCC CTC AGG GTT GTG AAC CGC AAG CGT CCA GCC
A   S   K   Y   Q   A   V   P   L   R   V   V   N   R   K   R   P   A

      2007       2016       2025       2034       2043       2052
CGG GAG GAT TGT CGG GGC CTC ACC GGC CCA CTG CAG AGC CTG GTC CCC AGT GCA
R   E   D   C   R   G   L   T   G   P   L   Q   S   L   V   P   S   A

      2061       2070       2079       2088       2097       2106
GAT GGC GAT GCT GAC AAC TGC TGT GTC CAG ATC ATG GGA GGC TAC TTC ACG TGG
D   G   D   A   D   N   C   C   V   Q   I   M   G   G   Y   F   T   W

      2115       2124       2133       2142       2151       2160
ACC CCA GAT GGA ATC CCC ACA CTG TCC AAC ATC ACC ATT CGT ATC CCC CGA GGC
T   P   D   G   I   P   T   L   S   N   I   T   I   R   I   P   R   G

      2169       2178       2187       2196       2205       2214
CAG CTG ACT ATG ATC GTG GGG CAG GTG GGC TGC GGC AAG TCC TCG CTC CTT CTA
Q   L   T   M   I   V   G   Q   V   G   C   G   K   S   S   L   L   L

      2223       2232       2241       2250       2259       2268
GCC GCA CTG GGG GAG ATG CAG AAG GTC TCA GGG GCT GTC TTC TGG AGC AGC AGC
A   A   L   G   E   M   Q   K   V   S   G   A   V   F   W   S   S   S
```

FIGURE 1F

```
      2277        2286        2295        2304        2313        2322
CTT CCT GAC AGC GAG ATA GGA GAG GAC CCC AGC CCA GAG CGG GAG ACA GCG ACC
L   P   D   S   E   I   G   E   D   P   S   P   E   R   E   T   A   T

      2331        2340        2349        2358        2367        2376
GAC TTG GAT ATC AGG AAG AGA GGC CCC GTG GCC TAT GCT TCG CAG AAA CCA TGG
D   L   D   I   R   K   R   G   P   V   A   Y   A   S   Q   K   P   W

      2385        2394        2403        2412        2421        2430
CTG CTA AAT GCC ACT GTG GAG GAG AAC ATC ATC TTT GAG AGT CCC TTC AAC AAA
L   L   N   A   T   V   E   E   N   I   I   F   E   S   P   F   N   K

      2439        2448        2457        2466        2475        2484
CAA CGG TAC AAG ATG GTC ATT GAA GCC TGC TCT CTG CAG CCA GAC ATC GAC ATC
Q   R   Y   K   M   V   I   E   A   C   S   L   Q   P   D   I   D   I

      2493        2502        2511        2520        2529        2538
CTG CCC CAT GGA GAC CAG ACC CAG ATT GGG GAA CGG GGC ATC AAC CTG TCT GGT
L   P   H   G   D   Q   T   Q   I   G   E   R   G   I   N   L   S   G

      2547        2556        2565        2574        2583        2592
GGT CAA CGC CAG CGA ATC AGT GTG GCC CGA GCC CTC TAC CAG CAC GCC AAC GTT
G   Q   R   Q   R   I   S   V   A   R   A   L   Y   Q   H   A   N   V

      2601        2610        2619        2628        2637        2646
GTC TTC TTG GAT GAC CCC TTC TCA GCT CTG GAT ATC CAT CTG AGT GAC CAC TTA
V   F   L   D   D   P   F   S   A   L   D   I   H   L   S   D   H   L
```

FIGURE 1G

```
        2655        2664        2673        2682        2691        2700
ATG CAG GCC GGC ATC CTT GAG CTG CTC CGG GAC GAC AAG AGG ACA GTG GTC TTA
M   Q   A   G   I   L   E   L   L   R   D   D   K   R   T   V   V   L

        2709        2718        2727        2736        2745        2754
GTG ACC CAC AAG CTA CAG TAC CTG CCC CAT GCA GAC TGG ATC ATT GCC ATG AAG
V   T   H   K   L   Q   Y   L   P   H   A   D   W   I   I   A   M   K

        2763        2772        2781        2790        2799        2808
GAT GGC ACC ATC CAG AGG GAG GGT ACC TCA AGG ACT TCC AGA GGT CTG AAT GCC
D   G   T   I   Q   R   E   G   T   S   R   T   S   R   G   L   N   A

        2817        2826        2835        2844        2853        2862
AGC TCT TTG AGC ACT GGA AGA CCT CAT GAA CCG ACA GGA CCA AGA GCT GGA GAA
S   S   L   S   T   G   R   P   H   E   P   T   G   P   R   A   G   E

        2871        2880        2889        2898        2907        2916
GGA AAT GTC ACA GAG AGA AAA GCC ACA GAG CCA CCC AGG GCC TAT CTC GTG CCA
G   N   V   T   E   R   K   A   T   E   P   P   R   A   Y   L   V   P

        2925        2934        2943        2952        2961        2970
TGT CCT CGA AGG GAT GGC CTT CTG CAG GAT GAG GAA GAG GAG GAA GAG GAG GCA
C   P   R   R   D   G   L   L   Q   D   E   E   E   E   E   E   E   A

        2979        2988        2997        3006        3015        3024
GCT GAG AAC GAG GAG GAT GAC TAC CTG TCG TCC ATG CTG CAC CAG CGT GCT GAG
A   E   N   E   E   D   D   Y   L   S   S   M   L   H   Q   R   A   E
```

FIGURE 1H

```
        3033        3042        3051        3060        3069        3078
ATC CCA TGG CGA GCC TGC NCC AAG TAC CTG TCC TCC GCC GGC ATC CTG CTC CTG
I   P   W   R   A   C   X   K   Y   L   S   S   A   G   I   L   L   L

        3087        3096        3105        3114        3123        3132
TCG TTG CTG GTC TTC TCA CAG CTG CTC AAG CAC ATG GTC CTG GTG GCC ATC GAC
S   L   L   V   F   S   Q   L   L   K   H   M   V   L   V   A   I   D

        3141        3150        3159        3168        3177        3186
TAC TGG CTG GCC AAG TGG ACC GAC AGC GCC CTG ACC CTG ACC CCT GCA ACC AGG
Y   W   L   A   K   W   T   D   S   A   L   T   L   T   P   A   T   R

        3195        3204        3213        3222        3231        3240
AAC TGC TCC CTC AAC CAG GAG TGC ACC CTC AAC CAG ACT GTC TAT GCC TTG GTG
N   C   S   L   N   Q   E   C   T   L   N   Q   T   V   Y   A   L   V

        3249        3258        3267        3276        3285        3294
TTC ACG GTG CTC TGC AGC CTG GGC ATT GTG CTG TGC CTC GTC ACG TCT GTC ACT
F   T   V   L   C   S   L   G   I   V   L   C   L   V   T   S   V   T

        3303        3312        3321        3330        3339        3348
GTG GAG TGG ACA GGG CTG AAG GTG GCC AAG AGA CTG CAC CGC AGC CTG CTA AAC
V   E   W   T   G   L   K   V   A   K   R   L   H   R   S   L   L   N

        3357        3366        3375        3384        3393        3402
CGG ATC ATC CTA GCC CCC ATG AGG TTT TTT GAG ACC ACG CCC CTT GGG AGC ATC
R   I   I   L   A   P   M   R   F   F   E   T   T   P   L   G   S   I
```

FIGURE 11

```
        3411        3420        3429        3438        3447        3456
CTG AAC AGA TTT TCA TCT GAC TGT AAC ACC ATC GAC CAG CAC ATC CCA TCC ACG
L   N   R   F   S   S   D   C   N   T   I   D   Q   H   I   P   S   T

        3465        3474        3483        3492        3501        3510
CTG GAG TGC CTG AGC CGC TCC ACC CTG CTC TGT GTC TCA GCC CTG GCC GTC ATC
L   E   C   L   S   R   S   T   L   L   C   V   S   A   L   A   V   I

        3519        3528        3537        3546        3555        3564
TCC TAT GTC ACA CCT GTG TTC CTC GTG GCC CTC TTG CCC CTG GCC ATC GTG TGC
S   Y   V   T   P   V   F   L   V   A   L   L   P   L   A   I   V   C

        3573        3582        3591        3600        3609        3618
TAC TTC ATC CAG AAG TAC TTC CGG GTG GCG TCC AGG GAC CTG CAG CAG CTG GAT
Y   F   I   Q   K   Y   F   R   V   A   S   R   D   L   Q   Q   L   D

        3627        3636        3645        3654        3663        3672
GAC ACC ACC CAG CTT CCA CTT CTC TCA CAC TTT GCC GAA ACC GTA GAA GGA CTC
D   T   T   Q   L   P   L   L   S   H   F   A   E   T   V   E   G   L

        3681        3690        3699        3708        3717        3726
ACC ACC ATC CGG GCC TTC AGG TAT GAG GCC CGG TTC CAG CAG AAG CTT CTC GAA
T   T   I   R   A   F   R   Y   E   A   R   F   Q   Q   K   L   L   E

        3735        3744        3753        3762        3771        3780
TAC ACA GAC TCC AAC AAC ATT GCT TCC CTC TTC CTC ACA GCT GCC AAC AGA TGG
Y   T   D   S   N   N   I   A   S   L   F   L   T   A   A   N   R   W
```

FIGURE 1J

```
        3789        3798        3807        3816        3825        3834
CTG GAA GTC CGA ATG GAG TAC ATC GGT GCA TGT GTG GTG CTC ATC GCA GCG GTG
L   E   V   R   M   E   Y   I   G   A   C   V   V   L   I   A   A   V

        3843        3852        3861        3870        3879        3888
ACC TCC ATC TCC AAC TCC CTG CAC AGA GAG CTC TCT GCT GGC CTG GTG GGC CTG
T   S   I   S   N   S   L   H   R   E   L   S   A   G   L   V   G   L

        3897        3906        3915        3924        3933        3942
GGC CTT ACC TAC GCC CTA ATG GTC TCC AAC TAC CTC AAC TGG ATG GTG AGG AAC
G   L   T   Y   A   L   M   V   S   N   Y   L   N   W   M   V   R   N

        3951        3960        3969        3978        3987        3996
CTG GCA GAC ATG GAG CTC CAG CTG GGG GCT GTG AAG CGC ATC CAT GGG CTC CTG
L   A   D   M   E   L   Q   L   G   A   V   K   R   I   H   G   L   L

        4005        4014        4023        4032        4041        4050
AAA ACC GAG GCA GAG AGC TAC GAG GGG CTC CTG GCA CCA TCG CTG ATC CCA AAG
K   T   E   A   E   S   Y   E   G   L   L   A   P   S   L   I   P   K

        4059        4068        4077        4086        4095        4104
AAC TGG CCA GAC CAA GGG AAG ATC CAG ATC CAG AAC CTG AGC GTG CGC TAC GAC
N   W   P   D   Q   G   K   I   Q   I   Q   N   L   S   V   R   Y   D

        4113        4122        4131        4140        4149        4158
AGC TCC CTG AAG CCG GTG CTG AAG CAC GTC AAT GCC CTC ATC TCC CCT GGA CAG
S   S   L   K   P   V   L   K   H   V   N   A   L   I   S   P   G   Q
```

FIGURE 1K

```
        4167        4176        4185        4194        4203        4212
AAG ATC GGG ATC TGC GGC CGC ACC GGC AGT GGG AAG TCC TCC TTC TCT CTT GCC
K   I   G   I   C   G   R   T   G   S   G   K   S   S   F   S   L   A

        4221        4230        4239        4248        4257        4266
TTC TTC CGC ATG GTG GAC ACG TTC GAA GGG CAC ATC ATC ATT GAT GGC ATT GAC
F   F   R   M   V   D   T   F   E   G   H   I   I   I   D   G   I   D

        4275        4284        4293        4302        4311        4320
ATC GCC AAA CTG CCG CTG CAC ACC CTG CGC TCA CGC CTC TCC ATC ATC CTG CAG
I   A   K   L   P   L   H   T   L   R   S   R   L   S   I   I   L   Q

        4329        4338        4347        4356        4365        4374
GAC CCC GTC CTC TTC AGC GGC ACC ATC CGA TTT AAC CTG GAC CCT GAG AGG AAG
D   P   V   L   F   S   G   T   I   R   F   N   L   D   P   E   R   K

        4383        4392        4401        4410        4419        4428
TGC TCA GAT AGC ACA CTG TGG GAG GCC CTG GAA ATC GCC CAG CTG AAG CTG GTG
C   S   D   S   T   L   W   E   A   L   E   I   A   Q   L   K   L   V

        4437        4446        4455        4464        4473        4482
GTG AAG GCA CTG CCA GGA GGC CTC GAT GCC ATC ATC ACA GAA GGC GGG GAG AAT
V   K   A   L   P   G   G   L   D   A   I   I   T   E   G   G   E   N

        4491        4500        4509        4518        4527        4536
TTC AGC CAG GGA CAG AGG CAG CTG TTC TGC CTG GCC CGG GCC TTC GTG AGG AAG
F   S   Q   G   Q   R   Q   L   F   C   L   A   R   A   F   V   R   K
```

FIGURE 1L

```
        4545        4554        4563        4572        4581        4590
ACC AGC ATC TTC ATC ATG GAC GAG GCC ACG GCT TCC ATT GAC ATG GCC ACG GAA
T   S   I   F   I   M   D   E   A   T   A   S   I   D   M   A   T   E

        4599        4608        4617        4626        4635        4644
AAC ATC CTC CAA AAG GTG GTG ATG ACA GCC TTC GCA GAC CGC ACT GTG GTC ACC
N   I   L   Q   K   V   V   M   T   A   F   A   D   R   T   V   V   T

        4653        4662        4671        4680        4689        4698
ATC GCG CAT CGA GTG CAC ACC ATC CTG AGT GCA GAC CTG GTG ATC GTC CTG AAG
I   A   H   R   V   H   T   I   L   S   A   D   L   V   I   V   L   K

        4707        4716        4725        4734        4743        4752
CGG GGT GCC ATC CTT GAG TTC GAT AAG CCA GAG AAG CTG CTC AGC CGG AAG GAC
R   G   A   I   L   E   F   D   K   P   E   K   L   L   S   R   K   D

        4761        4770        4779        4788        4797        4806
AGC GTC TTC GCC TCC TTC GTC CGT GCA GAC AAG TGA CCT GCC AGA GCC CAA GTG
S   V   F   A   S   F   V   R   A   D   K

        4815        4824        4833        4842        4851        4860
CCA TCC CAC ATT CGG ACC CTG CCC ATA CCC CTG CCT GGG TTT TCT AAC TGT AAA

        4869        4878        4887        4896        4905        4914
TCA CTT GTA AAT AAA TAG ATT TGA TTA TTA AAA AAA AAA AAA AAA AAA AAA AAA

        4923        4932
AAA AAA AAA AAA AAA AAA A 3'
```

FIGURE 1M

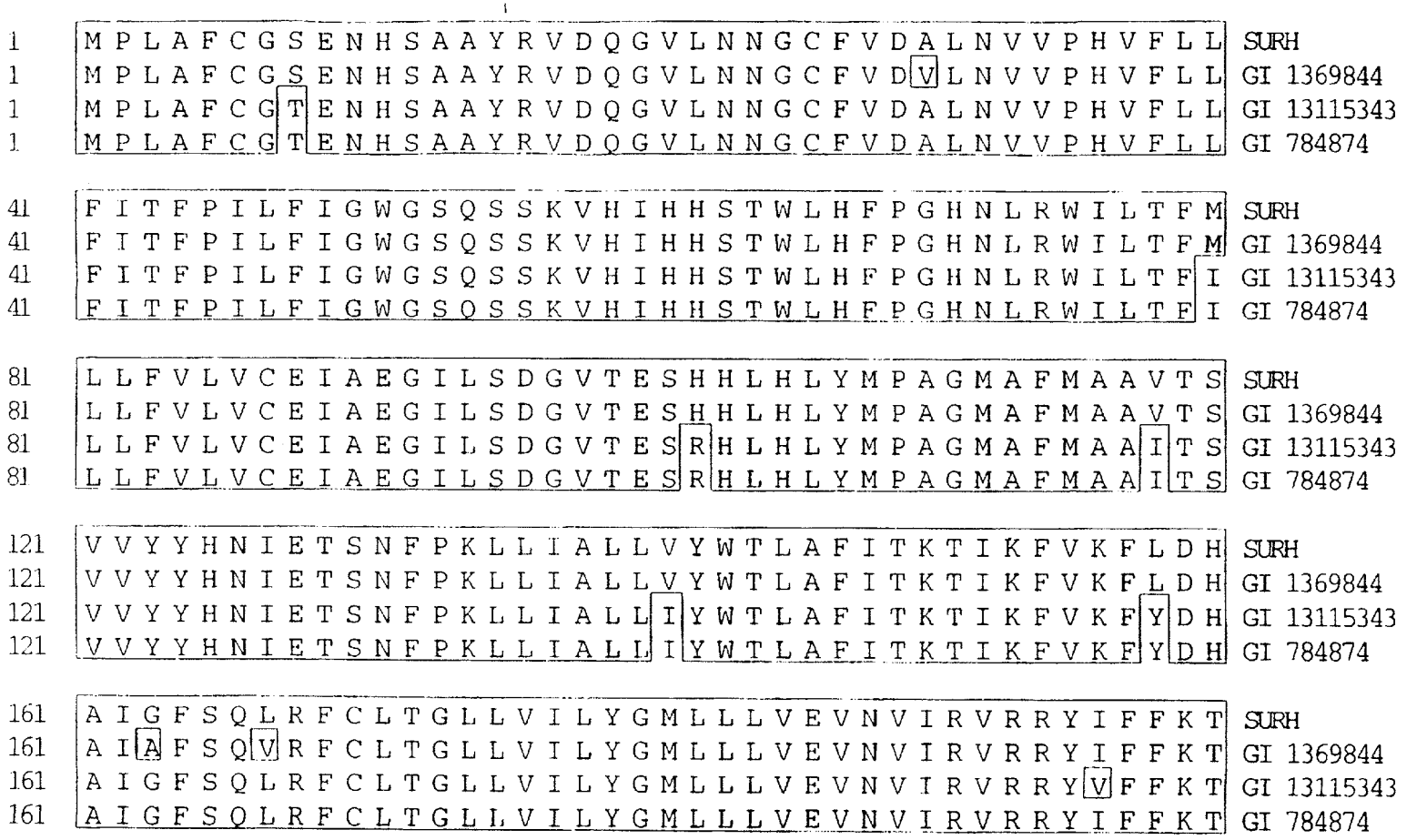

FIGURE 2A

```
201 PREVKPPEDLQDLGVRFLQPFVNLLSKGTYWWMNAFIKTA SURH
201 PREVKPPEDLQDLGVRFLQPFVNLLSKGTYWWMNAFIKTA GI 1369844
201 PREVKPPEDLQDLGVRFLQPFVNLLSKGTYWWMNAFIKTA GI 13115343
201 PREVKPPEDLQDLGVRFLQPFVNLLSKGTYWWMNAFIKTA GI 784874

241 HKKPXDLRAIGKLPIAMRALTNYQRLCEAFDAQ-RKDIQG SURH
241 HKKPIDLRAIGKLPIAMRALTNYQRLCEAFDAQVRKDIQG GI 1369844
241 HKKPIDLRAIGKLPIAMRALTNYQRLCLAFDAQARKDTQS GI 13115343
241 HKKPIDLRAIAKLPIAMRALTNYQRLCVAFDAQARKDTQS GI 784874

280 TQGARAIWQALSHAFGRRLVLSSTFRILADLLGFAGPLCI SURH
281 TQGARAIWQALSHAFGRRLVLSSTFRILADLLGFAGPLCI GI 1369844
281 QQGARAIWRALCHAFGRRLVLSSTFRILADLLGFAGPLCI GI 13115343
281 PQGARAIWRALCHAFGRRLILSSTFRILADLLGFAGPLCI GI 784874

320 FGIVDHLGKENDVFQPKTQFLGVYFVSSQEFLANAYVLAV SURH
321 FGIVDHLGKENDVFQPKTQFLGVYFVSSQEFLANAYVLAV GI 1369844
321 FGIVDHLGKENHVFQPKTQFLGVYFVSSQEFLGNAYVLAV GI 13115343
321 FGIVDHLGKENHVFQPKTQFLGVYFVSSQEFLGNAYVLAV GI 784874

360 LLFLALLLQRTFLQASYYVAIETGINLRGAIQTKIYNKIM SURH
361 LLFLALLLQRTFLQASYYVAIETGINLRGAIQTKIYNKIM GI 1369844
361 LLFLALLLQRTFLQASYYVAIETGINLRGAIQTKIYNKIM GI 13115343
361 LLFLALLLQRTFLQASYYVAIETGINLRGAIQTKIYNKIM GI 784874
```

FIGURE 2B

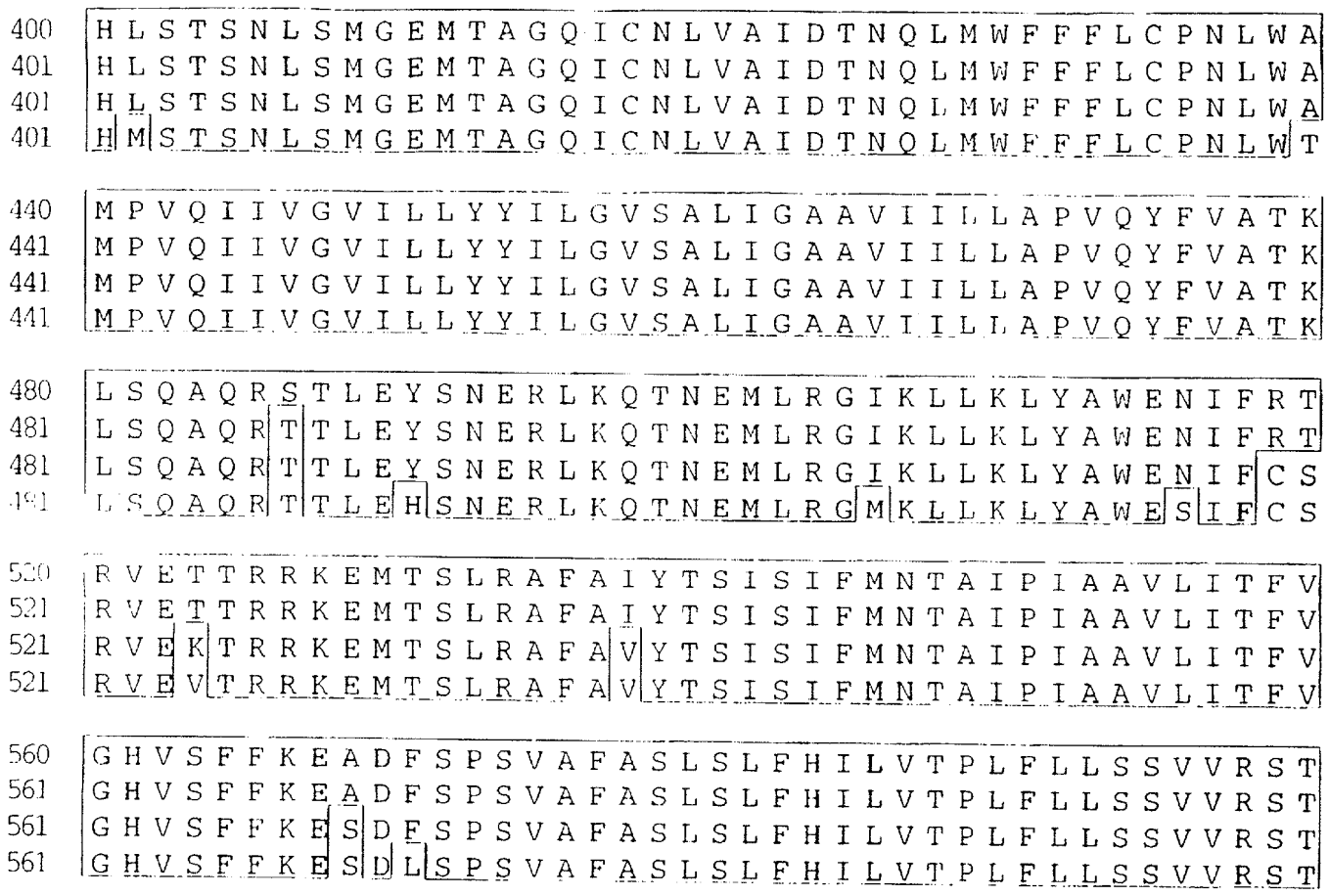

FIGURE 2C

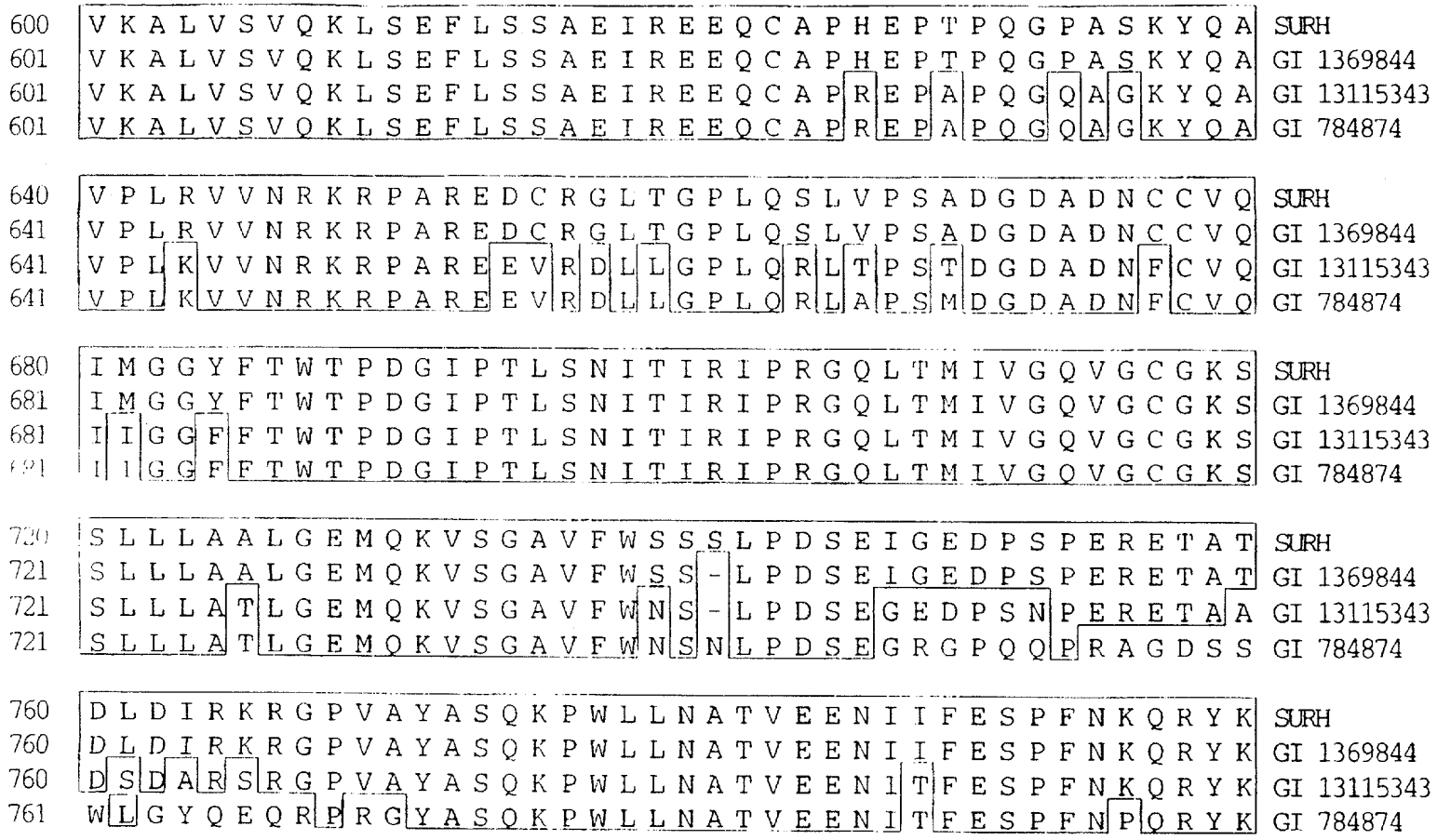

FIGURE 2D

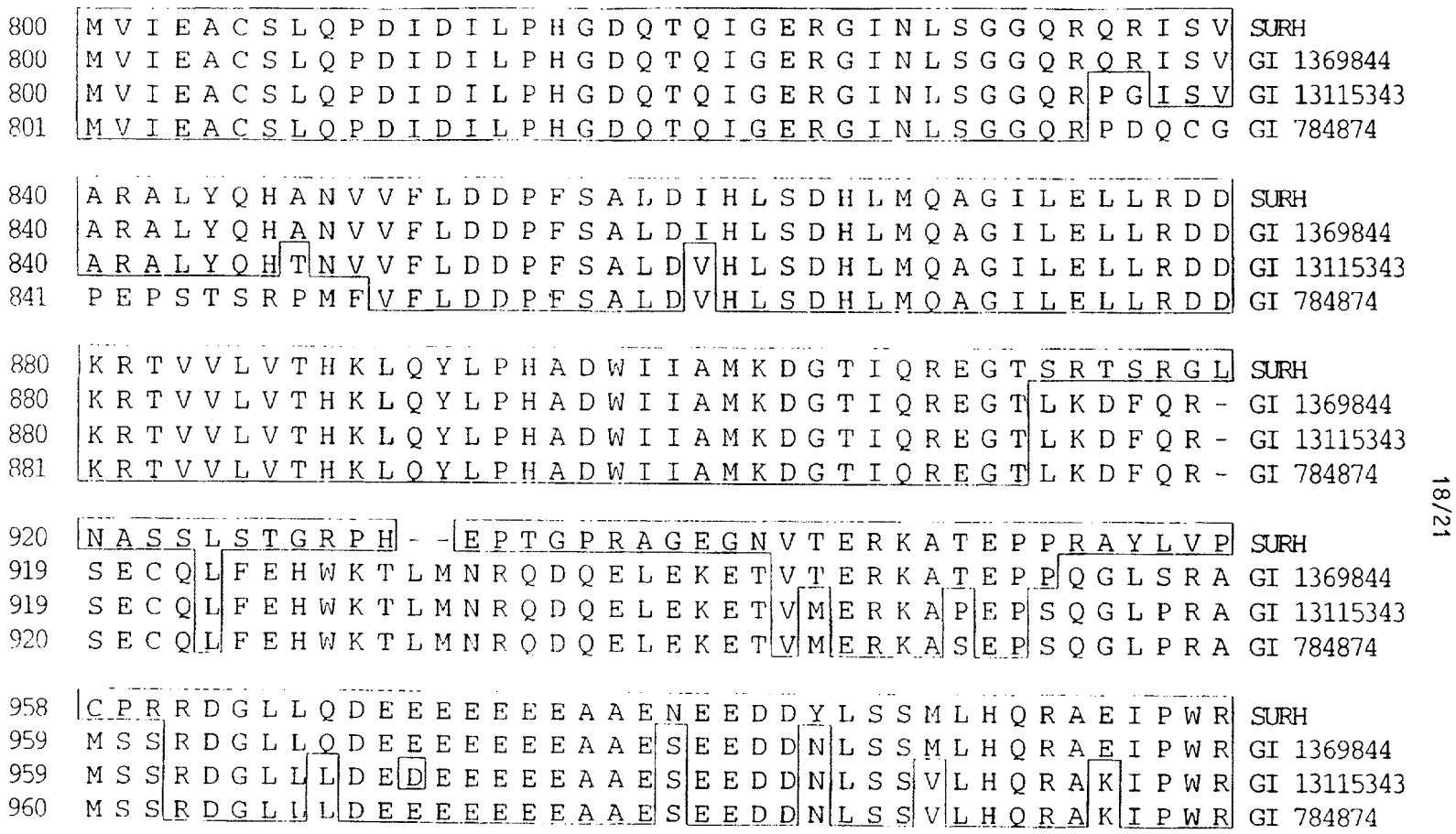

FIGURE 2E

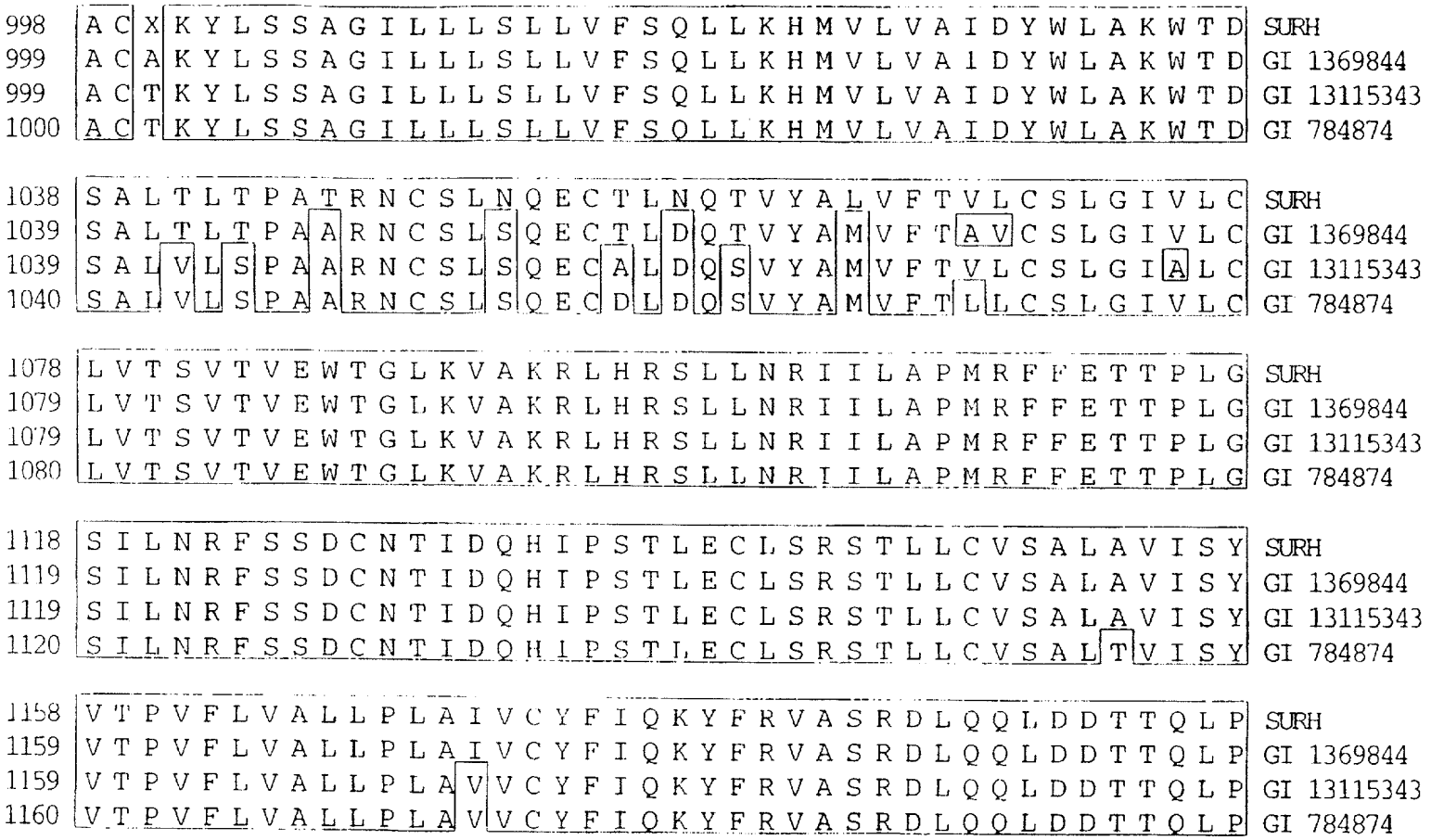

FIGURE 2F

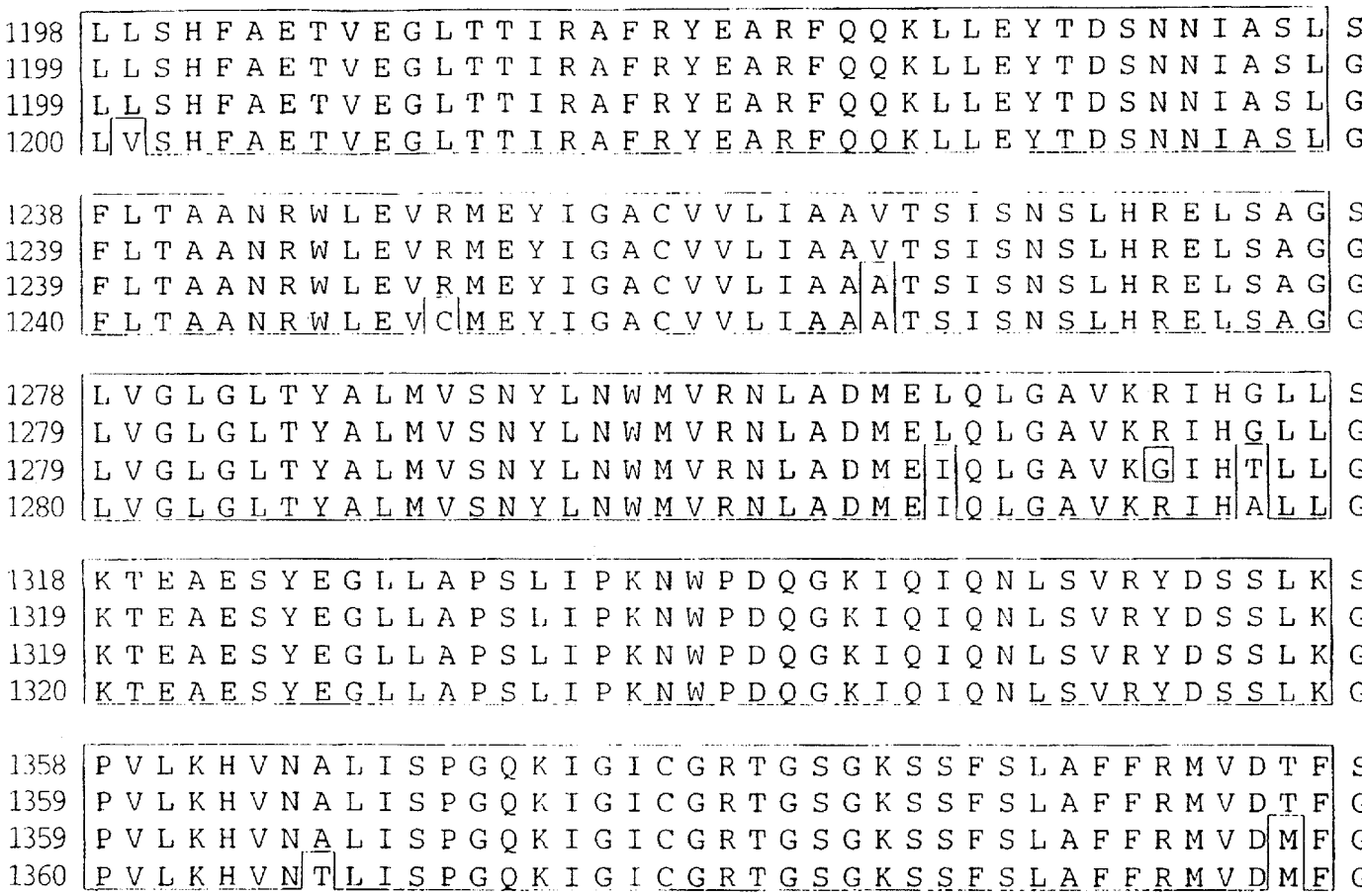

FIGURE 2G

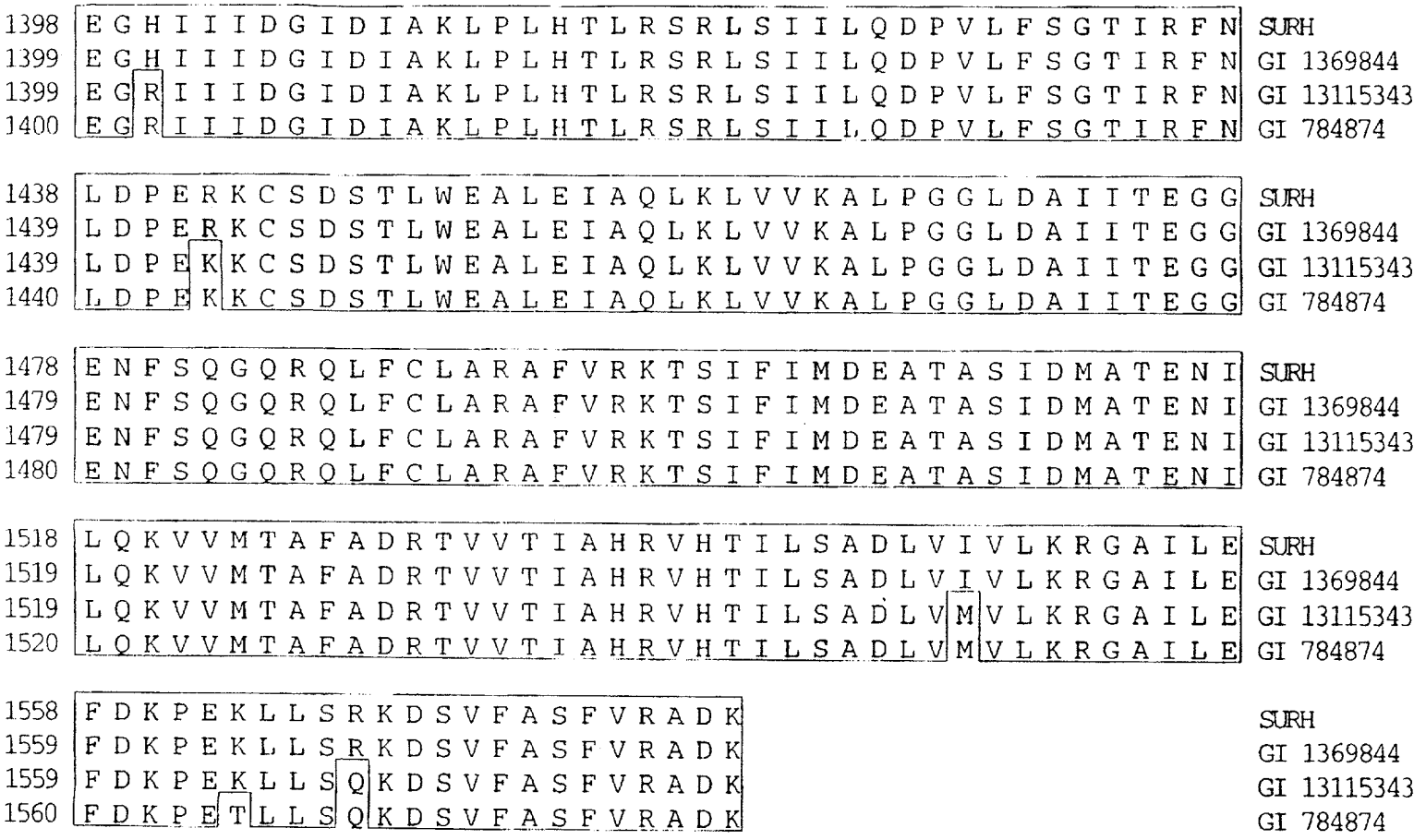
1398 E G H I I I D G I D I A K L P L H T L R S R L S I I L Q D P V L F S G T I R F N SURH
1399 E G H I I I D G I D I A K L P L H T L R S R L S I I L Q D P V L F S G T I R F N GI 1369844
1399 E G R I I I D G I D I A K L P L H T L R S R L S I I L Q D P V L F S G T I R F N GI 13115343
1400 E G R I I I D G I D I A K L P L H T L R S R L S I I L Q D P V L F S G T I R F N GI 784874
1438 L D P E R K C S D S T L W E A L E I A Q L K L V V K A L P G G L D A I I T E G G SURH
1439 L D P E R K C S D S T L W E A L E I A Q L K L V V K A L P G G L D A I I T E G G GI 1369844
1439 L D P E K K C S D S T L W E A L E I A Q L K L V V K A L P G G L D A I I T E G G GI 13115343
1440 L D P E K K C S D S T L W E A L E I A Q L K L V V K A L P G G L D A I I T E G G GI 784874
1478 E N F S Q G Q R Q L F C L A R A F V R K T S I F I M D E A T A S I D M A T E N I SURH
1479 E N F S Q G Q R Q L F C L A R A F V R K T S I F I M D E A T A S I D M A T E N I GI 1369844
1479 E N F S Q G Q R Q L F C L A R A F V R K T S I F I M D E A T A S I D M A T E N I GI 13115343
1480 E N F S Q G Q R Q L F C L A R A F V R K T S I F I M D E A T A S I D M A T E N I GI 784874
1518 L Q K V V M T A F A D R T V V T I A H R V H T I L S A D L V I V L K R G A I L E SURH
1519 L Q K V V M T A F A D R T V V T I A H R V H T I L S A D L V I V L K R G A I L E GI 1369844
1519 L Q K V V M T A F A D R T V V T I A H R V H T I L S A D L V M V L K R G A I L E GI 13115343
1520 L Q K V V M T A F A D R T V V T I A H R V H T I L S A D L V M V L K R G A I L E GI 784874
1558 F D K P E K L L S R K D S V F A S F V R A D K SURH
1559 F D K P E K L L S R K D S V F A S F V R A D K GI 1369844
1559 F D K P E K L L S Q K D S V F A S F V R A D K GI 13115343
1560 F D K P E T L L S Q K D S V F A S F V R A D K GI 784874

FIGURE 2H

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 97/17744

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/705 A61K38/17 C12Q1/68 G01N33/68
C07K16/28

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	WO 95 28411 A (BAYLOR COLLEGE MEDICINE :UNIV TEXAS (US)) 26 October 1995 see claims 1-50 ---	1-17
X	DATABASE EMBL HUMAN SEQUENCES EMBL Entry : Emhum2:Hssurlrna, 16 July 1996 GONZALEZ G., AGUILAR-BRYAN L., BRYAN J.:: "Human beta cell sulfonylurea receptor, SUR1, expression" XP002052209 see the whole document --- -/--	1-17

[X] Further documents are listed in the continuation of box C

[X] Patent family members are listed in annex

* Special categories of cited documents

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 January 1998	30/01/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Nauche, S

Form PCT/ISA/210 (second sheet) (July 1992)

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No
X	THOMAS PM ET AL: "Inactivation of the first nucleotide-binding fold of the sulfonylurea receptor, and familial persistent hyperinsulinemic hypoglycemia of infancy." AM J HUM GENET, SEP 1996, 59 (3) P510-8. UNITED STATES. XP002052206 see the whole document ---	1-17
X	AGUILAR-BRYAN L ET AL: "Cloning of the beta cell high-affinity sulfonylurea receptor: a regulator of insulin secretion." SCIENCE, APR 21 1995, 268 (5209) P423-6. UNITED STATES. XP002052207 cited in the application see the whole document ---	1-17
X	THOMAS, PAMELA L. ET AL.: "Mutations in the sulfonylurea recepetor gene in familial persistent hyperinsulinemic Hypoglycemia of Infancy" SCIENCE, APR 21 1995, 268 (5209) P426-9. UNITED STATES. XP002052208 see the whole document -----	1-17

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 97/17744

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9528411 A	26-10-95	AU 2285595 A	10-11-95
		CA 2187945 A	26-10-95
		EP 0789705 A	20-08-97

CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/705, A61K 38/17, C12Q 1/68, G01N 33/68, C07K 16/28

A1

(11) International Publication Number: WO 98/14571

(43) International Publication Date: 9 April 1998 (09.04.98)

(21) International Application Number: PCT/US97/17744

(22) International Filing Date: 2 October 1997 (02.10.97)

(30) Priority Data:
08/726,320 3 October 1996 (03.10.96) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 08/726,320 (CIP)
Filed on 3 October 1996 (03.10.96)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: AU-YOUNG, Janice [US/US]; 1419 Kains Avenue, Berkeley, CA 94702 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). COLEMAN, Roger [US/US]; 260 Mariposa #2, Mountain View, CA 94041 (US).

(74) Agent: BILLINGS, Lucy, J.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AT, AU, BR, CA, CH, CN, DE, DK, ES, FI, IL, JP, KR, MX, NO, NZ, RU, SE, SG, US, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HUMAN SULFONYLUREA RECEPTOR SURH

(57) Abstract

The present invention provides a human sulfonylurea receptor (SURH) and the polynucleotides which identify and encode SURH. The invention also provides genetically engineered expression vectors and host cells comprising the nucleic acid sequences encoding SURH and methods for producing the protein. The invention also provides pharmaceutical compositions containing SURH, agonists to SURH, or antagonists to SURH, and in the use of such compositions for the prevention or treatment of diseases associated with the expression of SURH. Additionally, the invention provides for the use of antisense molecules to polynucleotides encoding SURH for the treatment of diseases associated with the expression of SURH. The invention also provides diagnostic assays which utilize the polynucleotide, or fragments or the complement thereof, to hybridize to the genomic sequence or transcripts of polynucleotides encoding SURH, or anti-SURH antibodies which specifically bind to SURH.

*(Referred to in PCT Gazette No. 38/1998, Section II)
BNSDOCID: <WO___9814571A1_!A>

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		